Clonal evolution in chronic lymphocytic leukemia is scant in relapsed but accelerated in refractory cases after chemo(immune)therapy

Zapatka, Marc ; Tausch, Eugen ; Öztürk, Selcen ; Yosifov, Deyan Yordanov ; Seiffert, Martina ; Zenz, Thorsten ; Schneider, Christof ; Blöhdorn, Johannes ; Döhner, Hartmut ; Mertens, Daniel ; Lichter, Peter ; Stilgenbauer, Stephan

Abstract: Clonal evolution is involved in the progression of chronic lymphocytic leukemia (CLL). To link evolutionary patterns to different disease courses, we performed a long-term longitudinal mutation profiling study of CLL patients. Tracking somatic mutations and their changes in allele frequency over time and assessing the underlying cancer cell fraction revealed highly distinct evolutionary patterns. Surprisingly, in long-term stable disease and in relapse after long-lasting clinical response to treatment, clonal shifts are minor. In contrast, in refractory disease major clonal shifts occur although there is little impact on leukemia cell counts. As this striking pattern in refractory cases is not linked to a strong contribution of known CLL driver genes, the evolution is mostly driven by treatment-induced selection of sub-clones, underlining the need for novel, non-genotoxic treatment regimens.

DOI: https://doi.org/10.3324/haematol.2020.265777

haematologica
Journal of The Ferrata Storti Foundation

Clonal evolution in chronic lymphocytic leukemia is scant in relapsed but accelerated in refractory cases after chemo(immune)therapy

by Marc Zapatka, Eugen Tausch, Selcen Öztürk, Deyan Yordanov Yosifov, Martina Seiffert, Thorsten Zenz, Christof Schneider, Johannes Blöhdorn, Hartmut Döhner, Daniel Mertens, Peter Lichter, and Stephan Stilgenbauer

Haematologica 2021 [Epub ahead of print]

Citation: Marc Zapatka, Eugen Tausch, Selcen Öztürk, Deyan Yordanov Yosifov, Martina Seiffert, Thorsten Zenz, Christof Schneider, Johannes Blöhdorn, Hartmut Döhner, Daniel Mertens, Peter Lichter, and Stephan Stilgenbauer. Clonal evolution in chronic lymphocytic leukemia is scant in relapsed but accelerated in refractory cases after chemo(immune)therapy.
Haematologica. 2021; 106:xxx
doi:10.3324/haematol.2020.265777

Clonal evolution in chronic lymphocytic leukemia is scant in relapsed but accelerated in refractory cases after chemo(immune)therapy.

* Marc Zapatka [1]+, * Eugen Tausch [2]+, Selcen Öztürk [1], Deyan Yordanov Yosifov [2,3], Martina Seiffert [1], Thorsten Zenz [4], Christof Schneider [2], Johannes Blöhdorn [2], Hartmut Döhner [2], Daniel Mertens [2,3], # Peter Lichter [1], # Stephan Stilgenbauer [2]

* MZ and ET co-first authors
# PL and SS corresponding authors

1 Division of Molecular Genetics, German Cancer Research Center, Heidelberg, 69120, Germany
2 Department of Internal Medicine III, Ulm University Hospital Ulm, 89081, Germany
3 Mechanisms of Leukemogenesis, German Cancer Research Center (DKFZ), Heidelberg, 69120, Germany
4 University Hospital and University of Zürich, 8091, Switzerland

Clonal evolution in relapsed/refractory CLL cases

Peter Lichter, Molecular Genetics B060, Im Neuenheimer Feld 280, 69120 Heidelberg, Germany, phone +49-6221-42-4619, fax +49-6221-42-4639
Stephan Stilgenbauer, Department of Internal Medicine III, Ulm University, Albert-Einstein Allee 23, 89081 Ulm, Germany, phone +49-731-50045521, fax +49-731-50045525
Corresponding author:
Peter.Lichter@dkfz-heidelberg.de
Stephan.Stilgenbauer@uniklinik-ulm.de (Lead Contact)

Acknowledgements
We thank all patients and physicians, especially Andrea Schnaiter, for donating samples and participating in this study. We thank Michael Hain and Rolf Kabbe for computational support. We thank Stephan Wolf and the High Throughput Sequencing unit of the Genomics & Proteomics Core Facility, German Cancer Research Center (DKFZ), for providing excellent sequencing services.

Contributions
Conceptualization Ideas, PL SS; Software MZ; Methodology ET; Validation ET; Formal Analysis MZ, ET, DYY; Investigation MZ, ET, SÖ; Resources TZ, CS, JB, HD, ET, PL, SS; Data Curation MZ, ET, SS; Writing –Original Draft MZ, ET, SÖ, MS, DM, PL, SS; Writing –Review & Editing MZ, ET, SÖ, DYY, MS, TZ, CS, JB, DM, PL, SS; Visualization MZ, ET, SÖ, DYY, DM; Supervision PL, SS; Project Administration PL, SS; Funding Acquisition MZ, ET, TZ, PL, SS.

Disclosures
Daniel Mertens was supported by DFG (SFB1074 subproject B1 and B2) and ERA-NET "FIRE-CLL".
Marc Zapatka and Martina Seiffert were supported by ERA-NET "FIRE-CLL" and BMBF "PRECISE". Stephan Stilgenbauer received support from DFG (SFB1074 subproject B1 and B2, BMBF "PRECISE" and ERA-NET "FIRE CLL". Stephan Stilgenbauer received honoraria and research support from AbbVie, AstraZeneca, Celgene, Gilead, GSK, Hoffmann La-Roche, Janssen, Novartis. Eugen Tausch received honoraria from the Speakers Bureau, Advisory board and travel support of Hoffmann La-Roche and AbbVie.

Abstract

Clonal evolution is involved in the progression of chronic lymphocytic leukemia (CLL). To link evolutionary patterns to different disease courses, we performed a long-term longitudinal mutation profiling study of CLL patients. Tracking somatic mutations and their changes in allele frequency over time and assessing the underlying cancer cell fraction revealed highly distinct evolutionary patterns. Surprisingly, in long-term stable disease and in relapse after long-lasting clinical response to treatment, clonal shifts are minor. In contrast, in refractory disease major clonal shifts occur although there is little impact on leukemia cell counts. As this striking pattern in refractory cases is not linked to a strong contribution of known CLL driver genes, the evolution is mostly driven by treatment-induced selection of sub-clones, underlining the need for novel, non-genotoxic treatment regimens.

Introduction

Cancer can be conceptualized as an evolutionary process within a given organism [1, 2]. By increasing the fitness of cancer cells, mutations enable sub-clones to outcompete non-malignant cells and less adapted cancer cell clones. Furthermore, clonal evolution allows selection of cell populations that are resistant to therapy or responsible for disease recurrence. For some tumor entities like acute myeloid leukemia (AML), the concept of cancer initiating cells seems to account for tumor relapses without further genetic evolution. For other malignancies however, it is more likely that additional mutations play a crucial role in tumor recurrence. This is also true for chronic lymphocytic leukemia (CLL), where progression and clonal evolution has been analyzed in the context of treatment induced genetic changes [3, 4].

Clinically, CLL is characterized by a highly variable course. The survival time of patients varies between months and decades. Often patients remain untreated for many years until clinical symptoms require therapeutic intervention [5]. Despite high rates of initial treatment response, a

major clinical challenge is the occurrence of refractory disease that does not respond to treatment. Refractory cases are often characterized by a deletion and/or a mutation in the tumor suppressor gene *TP53* located on the short arm of chromosome 17 (del17p/*TP53*mut). Although a number of recurrently mutated genes were identified in CLL that are of prognostic relevance [4, 6-9] del17p/*TP53*mut remains the strongest adverse prognostic factor for progression-free and overall survival in CLL [4, 10, 11]. The incidence of mutated or deleted *TP53* is below 3% in Binet A stage CLL representing cases with good prognosis or in the pre-malignant monoclonal B-cell lymphocytosis (MBL) state, but increases to 12% at time of first treatment initiation, and to more than 37% in chemotherapy refractory cohorts [8, 12, 13]. Despite this increase in cases with mutated or deleted *TP53* at later disease stages, clonal evolution has been considered rare in CLL.

Early cytogenetic and molecular cytogenetic studies reported unequivocal evidence for occurrence of clonal evolution in CLL, albeit rare [14-17]. More recently, high-resolution microarray and next-generation-sequencing (NGS) based approaches were applied to track subclonal heterogeneity and clonal evolution in CLL. Based on a SNP micro-array analysis of pretreatment and relapsed samples from 42 patients, DNA copy number variations (CNVs) were reported that expand or newly occur at relapse [18]. The respective genomic regions contain candidate driver genes of relapse and/or chemotherapy resistance. Somatic mutation profiling of CLL by NGS revealed recurrent gene alterations [19] and confirmed molecular heterogeneity [20]. The comprehensive analysis of 149 CLL cases allowed to distinguish clonal (*MYD88*, trisomy 12, and del(13q)) and subclonal (*SF3B1* and *TP53*) driver mutations [20] and this order was validated by the same group in a huge clinical study [4]. While a considerable number of driver genes and recurrent genomic alterations were identified via WES analysis of a cumulative number of more than 1000 CLL patients, there are only few studies deciphering changes of drivers over the course of disease. Mutation profiling of 3 CLL patients over time indicated heterogeneous clonal evolution patterns [21]. By a similar approach, 10 of 12 CLL cases treated with chemotherapy were shown to undergo evolution of sub-clones with respective driver mutations (*SF3B1* and *TP53*), while this

was detected in only 1 of 6 cases that were not treated [22]. While one study reported that clonal composition remained stable at disease progression and relapse [23] another study referred that 13 of 28 sequentially sampled cases underwent genetic change of >20% with 9 of them (but none of the non-evolving cases) also display epigenetic evolution[24].

There was a number of deep sequencing studies focusing on a targeted panel for candidate genes in CLL providing evidence of clonal outgrowth over time i.e. of *TP53* after treatment[25-27]. Despite that, the major focus here was on untreated patient samples and the response to therapy was not considered as a predictor of evolution. In addition, targeted analysis of a restricted number of drivers can give an idea of clonal rigidity, but not show emergence and outgrowth of new subclones characterized by variants not covered with the panel. A similar approach considering aberrations in addition to known driver mutations deciphers the history of these alterations by integrating longitudinal and cross-sectional data in 70 patients [28]. While the distinction of evolutionary early and late events showed a similar pattern to Landau et al, again the association with patient outcome has not been addressed. The biggest whole exome sequenced cohort with sequential sampling in CLL included 59 patients from CLL8 with samples before and after relapse to FC/FCR showing changes of cell fractions characterized by specific drivers as well as linear versus branched evolution patterns in 57 of 59 cases [4]. However, this group consisted only of relapsed cases with a missing control of refractory and long-term untreated patients. Due to the fact, that again type and duration of response were not considered as parameters, a link between treatment, outcome and dynamic genomic changes in CLL is barely explored. Although a connection of response to therapy and dynamic genomic changes is plausible, it remains unclear how clonal evolution is linked to long-term stable, to relapsed or to refractory disease.

In order to elucidate the clonal evolution of CLL cell populations in the presence or absence of therapy, we performed a long-term longitudinal mutation profiling study of a multifarious cohort of CLL patients with a well annotated patient history.

Aberrant *TP53* dictates the clinical course of the disease, it is a key driver of acquired resistance and potentially supersedes other parameters. Therefore, we excluded patients with del17p or mutated *TP53* status at baseline as we presumed that these patients had acquired the most relevant evolution marker already. Samples were obtained at different time points before and after treatment in three different clinical groups: i) long-term untreated cases with stable disease and no need for treatment over at least 4 years, ii) relapsed cases with durable response to therapy of at least 2 years, and iii) refractory cases without response to treatment (SD, PD) or cases that progressed with requirement of a subsequent therapy within 1 year. Whole exome sequencing was performed and data were subsequently partially validated by targeted resequencing of identified mutations.

Methods

Sample collection

We compiled an inventory of CLL patient samples before and after treatment and sequenced tumor and non-tumor control DNA (25 patients and 54 tumor samples including 21 patients with baseline samples prior to any therapy). Our inclusion criteria were: (i) no del17p or mutated TP53 status at baseline (ii) patients fitting to any of the three groups (a) long-term untreated cases with stable disease and no need for treatment over at least 4 years, (b) relapsed cases with durable response to therapy of at least 2 years, and (c) refractory cases without response to treatment (SD, PD) or cases that progressed with requirement of a subsequent therapy within 1 year.

All patients gave informed consent according to the Helsinki Declaration. Sample acquisition for sequencing purposes was approved by a local ethical review committee (Ethikkommision Ulm University, ethik-kommission@uni-ulm.de, 17.06.2008, 96/08-UBB/se).

Peripheral blood mononuclear cell samples were enriched for tumor (CD19+) and normal CD19- cells using MACS microbead cell separation (Miltenyi Biotec, Bergisch Gladbach, Germany). Genomic DNA was isolated from unsorted and sorted CLL cells using All Prep Kit (Qiagen,

Hilden, Germany). Quality and quantity of the purified DNA were assessed with the Qubit dsDNA BR Assay Kit (Lifetech technologies, Carlsbad, CA).

Sequencing

Whole Exome Sequencing was performed on Illumina HiSeq 2000 machines. Exome libraries were created using the TruSeq Exome Library Prep Kit or Agilent SureSelect enrichment Human Exome V4 Kit according to manufacturer's protocols. Alignment and variant calling were performed as described in[29].

Allele frequency changes in patient groups

Per patient SNVs with genotype change were identified and differences in alternative allele frequency (aAF) calculated between consecutive time points. aAFs were clustered per patient into six clusters to give each time point equal weight regardless of the number of SNVs detected. Each change in aAF was grouped according to the status (untreated, relapse, refractory) at the second time point. Differences in the distribution of the allele frequency changes between the three patient groups were identified using a bootstrapped Kolmogorov-Smirnov test with n=10000.

CNV calling and calculation of absolute copy numbers

Estimation of the copy number state based on the exome sequencing data was achieved using Varscan 2 on the target regions[30]. Absolute copy numbers were calculated following[31].

Calculation of cancer cell fraction

Cancer cell fractions (CCFs) integrating sample purity (estimated by FACS), ploidy inferred from FISH, copy number states calculated from whole exome sequencing and allelic fraction and coverage of somatic variants were calculated for the patients with available germline samples following the approach outlined in [20].

Estimation of clonal composition by TrAP

Changes in clonal tumor composition were calculated integrating the CCF at the respective time points using TrAP [32].

Quantification of DNA methylation and estimation of correlation between time points

DNA methylation from the first and second time point of 10 patient phases (3 long-term untreated, 2 relapsed and 5 refractory) was assessed by Illumina Infinium HumanMethylation450 BeadChips according to the manufacturer's protocol.

Details on the individual approaches are further described in the supplementary methods.

Results

The clinical course of patients grouped into distinct phases

The clonal evolution in malignant B-cell populations of CLL patients was studied by longitudinal analyses in a total of 25 patients and 54 tumor samples. For 21 patients, the baseline sample was obtained prior to any CLL therapy, whereas 4 additional patients were pretreated before enrollment in our study. A common case history in CLL can consist of different phases including an untreated phase with a watch and wait strategy in the beginning followed by one or more therapies with either durable or very short responses or even refractoriness to the ongoing treatment. We observed such clinical phases in our patients throughout their individual medical history. For example, some of the long-term untreated patients required therapy at a later stage (e.g. HU-1-06) and some patients with initially long-lasting response became refractory after a subsequent treatment (e.g. HU-1-11). Therefore, we divided the individual patient histories into different clinical phases rather than using a rigid division of patients into categories. Individuals can go through several of these phases with sampling at the beginning and at the end of each phase. Three clinical disease patterns were distinguished and in total we identified 29 phases: 6 phases were evaluated as long-term untreated, 5 as relapsed after initially durable response to

therapy, and 18 as treatment refractory. Details of the clinical course of patients and patient phases including treatment, treatment response and sampling, as well as cytogenetic grouping and the IGHV mutation status are presented in supplementary tables 1-3 and in supplementary figure 1.

Increased mutation rate is associated with refractory disease

Identification of mutations was performed by comparative whole exome sequencing of CD19+ enriched peripheral blood mononuclear cells (PBMC) and, as non-malignant control, the sorted CD19-negative fraction of PBMC from the same patient. Over the course of this longitudinal study, no IGHV status switch was identified. In IGHV mutated cases, the major IGHV clone did not change, and IGHV mutations and SNP fingerprinting were used to confirm sample identity. Based on limited material for sorting of non-neoplastic cells, for 19 of 25 patients a non-tumor control was available for mutation detection. Applying established algorithms [33] for the calling of single nucleotide variations (SNVs) and small insertions and deletion (Indels), we observed an average of 15.1 mutations per sample (range 2-36)(supplementary table 4-6). A prediction of the response to therapy was not possible based on mutation numbers, as samples taken before long lasting response to therapy and before refractory disease had similar numbers of mutations (11.3 (range 1-30) and 15.8 (range 2-34) respectively *p*-value Mann-Whitney test p=0.36, figure 1A, B). Samples obtained before any therapy as well as post-therapeutic samples from relapsed patients had the lowest number with 13.5 (range 2-30) and 13.0 (range 6-25) mutations in contrast to refractory patients with 17.9 (range 4-36) mutations, respectively (figure 1B) (Kruskal Wallis test *p*-value= 0.30). We identified 1.5 known driver events per sample with the largest variation and highest number of SNVs/Indels in refractory CLL samples. All cases except HU-1-08, HU-1-11, and HU-1-21 harbored SNVs/Indels in known or candidate CLL driver genes [4, 34]. Indeed, candidates previously associated with adverse outcome like *BIRC3*, *EGR2* and *SAMHD* were identified predominantly in refractory cases, but preceded good response to (chemo)therapy

and therefore did not determine outcome (e.g., patients HU-1-19 or HU-1-15). In addition, this study revealed genes that were so far not associated with CLL but were mutated in more than one of the analyzed patients: MC5R, MYH2, RFX7, ROBO2 and SLITRK5.

Clonal evolution of leukemic cells is dominant in patients with refractory disease

Clonal evolution was modeled on the basis of SNVs that were assessed in longitudinal sample collections. FISH analysis with a panel of diagnostic probes [10] in a subset of samples revealed near diploidy of the neoplastic cells. Interestingly, no changes in cytogenetic aberrations in long-term untreated phases could be identified based on FISH data (supplementary table 1). Most patients retained their karyotype after treatment, but HU-1-19 acquired a deletion in chromosome 17p. Since neoplastic B-cell content was generally higher than 80%, allele frequencies were used as basis for modeling evolution over time. To this aim, SNVs were identified that displayed variable allele frequencies (AFs) between the time points of molecular analysis. During long-term untreated phases, AFs remained stable, which is in accordance with an unchanged clonal composition. Although one might expect in tumors with relapse after therapy the occurrence of clonal evolution with the acquisition of new variants, we observed the opposite: such clinical phases show the same mutational landscapes both at baseline and at relapse, and major shifts in AFs occurred only exceptionally (figure 2). In sharp contrast, during phases of therapy refractoriness, we found dramatic alterations in clonal composition (figure 2, figure 3). Notably, in refractory phases these increases and reductions in AFs occurred within relatively short time intervals (median phase length: refractory 707 days, relapse 2395 days, untreated 2088 days, time span refractory phases vs. time span untreated, treated and relapsed Mann-Whitney test p= 0.00014, supplementary table 1), that were particularly much shorter than the phases in stable or relapsed cases. The high degree of allele frequency changes and the short time window over which these changes occurred indicate marked dynamics in the clonal shift, often notable in

tumors that appeared clinically unaffected by therapy (i.e. without remission and subsequent regrowth). These clonal shifts clearly indicate a change in the clonal composition, and strikingly they occur mostly during refractory phases, i.e. during treatment that does not successfully affect the clinical outcome.

We further quantified overall allele frequency changes in the three different groups of clinical phases independent of the further course of disease (figure 3, changes in AFs over time provided in supplementary figure 2). Clearly, a substantially higher variation of AFs is seen in the samples that reach the therapy refractory phase. Comparison of 28 subsequent time points in 25 patients identified significant differences in the allele frequency changes over time between different types of clinical phases, which indicates that the degree of change in the clonal composition is different in the three clinical groups (Kruskal-Wallis test $p=0.00262$, corrected based on permutation of phase labels). Furthermore, the untreated phases showed a significantly lower allele frequency change over time independent of the extent of time between sampling ($p<0.01$ for untreated vs relapsed and untreated vs relapsed/refractory; supplementary table 7).

Clonal expansion or reduction of individual sub-clones over time and with treatment

Integrating tumor purity, copy number state (supplementary figure 3) and allele frequencies, we inferred the cancer cell fractions (CCF, supplementary table 8) affected by individual mutations as described before [4]. In some patients and phases, known cancer drivers listed in the COSMIC mutation database could be linked to the changes in cancer cell fraction (supplementary figure 4, gene symbols from COSMIC highlighted in purple). For example, HU-1-13 showed a mutation in the cancer driver *EGFR* only in the untreated and refractory sample (CCF 7% coverage 33 and CCF 66%, coverage 57) (supplementary figure 4). The mutation was undetected (coverage 48) at relapse indicating a reduction of this clone below the detection limit at relapse. Furthermore, the fraction of cells carrying an *ANO1* mutation steadily increased from 0% over 2.3% to 22.0% in the refractory sample (coverage 48, 42 and 41). Interestingly, the major clone present at the

relapse and characterized by an *NLRP13* mutation (CCF 28.5%, coverage 21) was not detected any more in the refractory sample (coverage 25), indicating that this clone was lost during treatment or during progress.

HU-1-19 displayed similar shifts albeit with a different clonal composition, but also with elimination of a clone after treatment. The EGR2 variant changed from 2.1% to 27% mutant allele frequency (CCF 4%, coverage 47 and CCF 55%, coverage 37) during the treatment-free interval, but dropped below the detection limit after first treatment (coverage 56). In addition, the major clone at the first time point characterized by MARK2 (CCF 37%, coverage 74) without treatment ("untreated") was slightly less prominent at the second time point without treatment (CCF 34%, coverage 54) and undetectable after treatment (coverage 60). These observations indicate a gradual change in clonal composition during an untreated phase of six years and a significant clonal replacement after treatment.

Refractory CLL is associated with a branched evolution of leukemic cells

In order to group observed clonal changes into different patterns of evolution, we analyzed overall AF changes between all possible pairs of consecutive samples that we grouped into disease phases for different types of evolution. On the basis of time-dependent changes of CCF, when significant AF changes were unidirectional, these evolution patterns were classified as co-evolution (also termed "linear evolution") [22]. In contrast, evolution was classified as "branched" when different significant changes concomitantly increased and decreased in CCF between consecutive samplings and were thus "bidirectional".

As depicted in figure 4A, we observed in the group of long-term untreated patients phases in the clinical course where substantial shifts in the clonal composition occurred only after treatment. These shifts resulted either in co-evolution of few sub-clones while other sub-clones were lost (e.g. HU-1-15), or a simultaneous decrease and increase of different subclones indicative of branched evolution. Long-term untreated phases display mainly co-evolution patterns (4 of 5),

while only one patient (HU-1-19), who required treatment after six years, followed a branched evolution pattern and became refractory to a subsequent treatment. In phases preceding relapse, co-evolution appeared less frequent (3 of 5). Instead, relapsed and refractory CLL showed a tendency towards more frequent branched evolution than untreated CLL (6 of 10 vs. 1 of 5, Fisher exact test p=0.28), a pattern that was e.g. observed in the relapsed and treatment refractory phases of a single patient (HU-1-13). Interestingly, this patient was treated with FCR (Fludarabine, Cyclophosphamide, Rituximab) in 2005 and in 2009 again achieving a complete response (CR) each time, but at the second time with shorter duration. The differences in evolution types shows a trend towards more branched evolution in relapsed and even more in refractory phases (figure 4B). To assess the dynamic on a cellular level we inferred the clonal composition based on the cancer cell fractions using TrAP (figure 5). Interestingly the major subclone at refractory time point in HU-1-19 (clone 3, 41.5% clone fraction) is already present to a minor extend at the second untreated time point (clone fraction 2.0%) whereas the major clone at the first time point further evolved gaining an additional set of mutations subdividing into clones 4, 5 and 6 (clone fraction of 13.7%, 21.3% and 10.5%). In contrast a clone (19.0% clone fraction) present at the second untreated time point defined by an EGR2 mutation is undetected at the other last time point. To confirm the evolutionary changes with an additional method, we performed epigenetic analysis of 20 samples corresponding to 10 phases. In line with the genetic data, large-scale evolution of methylation patterns was not present in any of the evaluated long-term untreated (n = 3) and relapsed (n = 2) phases displaying clonal changes of linear type while 3 of the 5 examined refractory phases featured profound changes in DNA methylation (supplementary figure 5). In all patients, even in the ones with few methylation changes, hypermethylation was concentrated in poised promoters and Polycomb-repressed regions, whereas hypomethylation occurred mostly in heterochromatin (data not shown; assignment of chromatin states was according to the published reference epigenome of CLL[35]). In spite of this

common pattern, we could not identify any specific CpGs that consistently changed methylation status throughout different patients or phases.

In summary, based on the patients analyzed here, the clonal evolution pattern seems to be linked to the disease phases, and increased changes in allele frequency and a branched evolution are significantly more frequent after treatment compared to untreated patient phases (figure 6).

Discussion

Medical history and disease course of patients with CLL is very individual. In this study, we examined whole exome sequencing data of CLL patients acquired at several time points during their disease and treatment course. Comparing consecutive samples from individual patients, we identified somatic mutations that were present in the leukemia cells and tracked over time the changes in allele frequency of these mutations and the underlying fraction of cancer cells that carried the respective mutations. By modelling the clonal composition using the software TrAP [32], we discovered different clonal evolution patterns and disease progression courses that were linked to the treatment and response history of the patients (figure 5). From the mutations and clonal changes that occur during CLL disease progression, we draw the following conclusions with respect to groups of genes, but also more conceptually with respect to clonal composition and evolution over time.

Recurrent mutations in genes were linked to CLL relapse in three different time and treatment dependent patterns [36]. First, one subset of genes initially displays subclonal mutations that are enriched after therapy. In contrast, mutations in a second set of genes remained clonally stable upon relapse. Finally, mutations in a third set of genes that are stable in most patients show clonal enrichment only in rare cases. However, these groups of mutations were not linked to a clinical phenotype. Furthermore, exponential-like growth patterns were recently associated with a larger

number of CLL drivers and short time to first treatment [37]. Of note, these data were derived from untreated CLL patients followed over time. In our patient cohort under the selective pressure of treatment, neither common genetic risk factors like IGHV or recurrent aberrations, nor variants or typically affected pathways are characteristic for a specific clinical course. And although the number of mutations increased slightly after treatment, this did not reflect or even predict outcome after therapy, nor did the number of (sub-)clones. Furthermore, clonal evolution was associated with treatment and indeed branched evolution was found more often in refractory cases, but not exclusively. These results reflect published data for relapsed cases after FCR therapy, which could also not link progression free survival to an evolution pattern after FC(R) therapy [4, 22]. Dividing our patient groups in long term responder and refractory cases allowed us in contrast to prior attempts to match the duration of response to the extent of the clonal shift. Counterintuitively, clonal evolution that was mostly dynamic occurred primarily in patients who displayed refractory disease, i.e. where major changes in clonal evolution happened under the guise of a clinically stable or progressing disease. Therefore, what correlated most with the duration of response to treatment was a highly dynamic evolutionary change among sub-clones, and this change was directly associated with refractory disease. In contrast and unexpectedly, relapse after initially durable response occurred mostly with the same sub-clones. We identified three distinctly different courses of clonal evolution that occurred under distinctly different treatment and response patterns. In refractory cases, clonal composition changed dramatically upon treatment failure and in patients 2, 4, and 18 this happened within only 3 months of therapy. Furthermore, in refractory phases, change in clonal composition was often accompanied by a profound shift in the bulk DNA methylation profile of the tumor, most probably reflecting different methylation profiles of the competing clones rather than *de novo* methylation changes, as it was previously shown that established CLL clones are epigenetically stable and changes in DNA methylation are unlikely to occur without genetic evolution [24]. As an example, patient HU-1-23 did not gain any new mutations between the two timepoints of his refractory phase but

underwent selection of particular pre-existent CLL clones according to the branched genetic evolution model and this was also manifested by a shift in DNA methylation of the bulk tumor. For the clinician managing the patient, “hidden“ selection of a resistant clone is masked by a tumor with a seemingly stable clinical phenotype, i.e. with a persistent lymphadenopathy and leukocytosis. This dynamic clonal change suggests either an increased evolutionary capacity in these patients or simply the presence of at least one resilient clone. Importantly a selective pressure of therapy is necessary to induce or catalyze this clonal change as provided by the clinical course of patients 15 and 19. Both have a longterm untreated phase without marked evolution but a strong shift after becoming refractory. Therefore, neither the underlying risk factors, the natural disease biology nor the type of evolution, which is branched in untreated and refractory phase in both patients, reflects the clinical course while the extent of evolution does. Importantly, we could show in addition that this process occurs also independently of *TP53* mutations, i.e. in a cohort of CLL patients without *TP53* aberrations before treatment. Thus, in these patients the clonal evolution is driven mostly by treatment that seems to select for resistant CLL clones.

Thus, our key finding is the striking observation of a clonal turnover during therapy in those patients, who were considered treatment refractory and therefore are assumed to have a stable tumor load.

In contrast, long-term untreated cases and late relapses are genomically stable, although they are observed over a much longer period of time. In the latter we found a remarkably stable genomic landscape considering that these patients received a therapy with a subsequent re-growing after a prolonged treatment-free interval. This stability is completely different to a tumor that is refractory and apparently unaffected by therapy, but in contrast displays a dramatic change in clonal composition. This opposing clinical and genomic phenotype at first appears counterintuitive. However, these different courses of clonal dynamics in relapsing and refractory patient phases could be explained by the preexistence of a resistant clone that after removal of the

bulk tumor by a treatment intervention will quickly grow out and fill the empty niche. If such resistant clones are absent, competition and outgrowth over time is still possible, so that the tumor regrows with an almost identical clonal composition. This finding mechanistically explains and underlines the relevance of the widely used clinical paradigm of repeating the previous treatment regimen when a good and long-lasting response is achieved: based on the same clonal composition at relapse, the clinician can expect another good response of the tumor to the treatment as the tumor has the same clonal composition as before the treatment. Interestingly, new treatment modalities like venetoclax may behave similarly due to the strong reduction of tumor load, comparable to chemotherapy: While patients treated with short and effective venetoclax containing combination therapies lack *BCL2* mutations at relapse, refractoriness to a long lasting venetoclax treatment associates with the outgrowth of a *BCL2* mutated clone displaying the same clonal shift towards drug resistance, that we observe here in chemotherapy refractory cases [38, 39]. On the other hand, ibrutinib may cause a more decelerated clonal shift due to its slow debulking treatment effect and also only slowly emerging resistant clones (i.e. point mutations in *BTK*/*PLCG2* [40]).

In summary despite the small patient cohort (n=25 with 54 time points), we identified a link between changes in the variant allele frequency and changes in clonal architecture, both of which are linked with shortened time to further treatment, i.e. treatment resistance. We found clonal evolution to occur without strong contribution of known CLL driver genes. However, there is a dramatic difference in clonal evolution patterns between relapsed and refractory samples, which highlights the importance of the treatment-induced clonal changes in relation to treatment response. This intrinsic characteristic of CLL evolution underlines the relevance of comparing the benefits of treatment compared to the watch-and-wait strategy that has a very low clonal evolution rate. Furthermore, the substantial clonal evolution in refractory disease highlights the

need for novel, non-genotoxic treatment regimens with targeted therapy that are less likely to induce clinical disease resistance by selecting out preexistent refractory sub-clones.

Data availability

Sequencing data have been deposited at the European Genome-Phenome Archive hosted at the EBI under accession EGAS00001003652. Data for the methylation arrays are accessible at GSE143411.

References

1. McGranahan N, Rosenthal R, Hiley CT, et al. Allele-Specific HLA Loss and Immune Escape in Lung Cancer Evolution. Cell. 2017;171(6):1259-1271.
2. Nowell PC. The clonal evolution of tumor cell populations. Science. 1976;194(4260):23-28.
3. Landau DA, Sun C, Rosebrock D, et al. The evolutionary landscape of chronic lymphocytic leukemia treated with ibrutinib targeted therapy. Nat Commun. 2017;8(1):2185.
4. Landau DA, Tausch E, Taylor-Weiner AN, et al. Mutations driving CLL and their evolution in progression and relapse. Nature. 2015;526(7574):525-530.
5. Hallek M, Shanafelt TD, Eichhorst B. Chronic lymphocytic leukaemia. Lancet. 2018;391(10129):1524-1537.
6. Rossi D, Rasi S, Fabbri G, et al. Mutations of NOTCH1 are an independent predictor of survival in chronic lymphocytic leukemia. Blood. 2012;119(2):521-529.
7. Rose-Zerilli MJJ, Forster J, Parker H, et al. The Correlation Between Deletion Architecture, ATM Mutational Status and BIRC3 Disruption in 11q-Deleted CLL. Blood. 2012;120(21):658-658.
8. Stilgenbauer S, Schnaiter A, Paschka P, et al. Gene mutations and treatment outcome in chronic lymphocytic leukemia: results from the CLL8 trial. Blood. 2014;123(21):3247-3254.
9. Puente XS, Bea S, Valdes-Mas R, et al. Non-coding recurrent mutations in chronic lymphocytic leukaemia. Nature. 2015;526(7574):519-524.
10. Dohner H, Stilgenbauer S, Benner A, et al. Genomic aberrations and survival in chronic lymphocytic leukemia. N Engl J Med. 2000;343(26):1910-1916.
11. Zenz T, Eichhorst B, Busch R, et al. TP53 mutation and survival in chronic lymphocytic leukemia. J Clin Oncol. 2010;28(29):4473-4479.
12. Schnaiter A, Paschka P, Rossi M, et al. NOTCH1, SF3B1, and TP53 mutations in fludarabine-refractory CLL patients treated with alemtuzumab: results from the CLL2H trial of the GCLLSG. Blood. 2013;122(7):1266-1270.
13. Winkelmann N, Rose-Zerilli M, Forster J, et al. Low frequency mutations independently predict poor treatment-free survival in early stage chronic lymphocytic leukemia and monoclonal B-cell lymphocytosis. Haematologica. 2015;100(6):e237-239.
14. Han T, Ohtaki K, Sadamori N, et al. Cytogenetic evidence for clonal evolution in B-cell chronic lymphocytic leukemia. Cancer Genet Cytogenet. 1986;23(4):321-328.
15. Oscier D, Fitchett M, Herbert T, Lambert R. Karyotypic evolution in B-cell chronic lymphocytic leukaemia. Genes Chromosomes Cancer. 1991;3(1):16-20.
16. Shanafelt TD, Witzig TE, Fink SR, et al. Prospective evaluation of clonal evolution during long-term follow-up of patients with untreated early-stage chronic lymphocytic leukemia. J Clin Oncol. 2006;24(28):4634-4641.

17. Stilgenbauer S, Sander S, Bullinger L, et al. Clonal evolution in chronic lymphocytic leukemia: acquisition of high-risk genomic aberrations associated with unmutated VH, resistance to therapy, and short survival. Haematologica. 2007;92(9):1242-1245.
18. Knight SJ, Yau C, Clifford R, et al. Quantification of subclonal distributions of recurrent genomic aberrations in paired pre-treatment and relapse samples from patients with B-cell chronic lymphocytic leukemia. Leukemia. 2012;26(7):1564-1575.
19. Puente XS, Pinyol M, Quesada V, et al. Whole-genome sequencing identifies recurrent mutations in chronic lymphocytic leukaemia. Nature. 2011;475(7354):101-105.
20. Landau DA, Wu CJ. Chronic lymphocytic leukemia: molecular heterogeneity revealed by high-throughput genomics. Genome Med. 2013;5(5):47.
21. Schuh A, Becq J, Humphray S, et al. Monitoring chronic lymphocytic leukemia progression by whole genome sequencing reveals heterogeneous clonal evolution patterns. Blood. 2012;120(20):4191-4196.
22. Landau DA, Carter SL, Stojanov P, et al. Evolution and impact of subclonal mutations in chronic lymphocytic leukemia. Cell. 2013;152(4):714-726.
23. Ojha J, Ayres J, Secreto C, et al. Deep sequencing identifies genetic heterogeneity and recurrent convergent evolution in chronic lymphocytic leukemia. Blood. 2015;125(3):492-498.
24. Oakes CC, Claus R, Gu L, et al. Evolution of DNA methylation is linked to genetic aberrations in chronic lymphocytic leukemia. Cancer Discov. 2014;4(3):348-361.
25. Rasi S, Khiabanian H, Ciardullo C, et al. Clinical impact of small subclones harboring NOTCH1, SF3B1 or BIRC3 mutations in chronic lymphocytic leukemia. Haematologica, 2016;101(4):135-138.
26. Nadeu F, Delgado J, Royo C, et al. Clinical impact of clonal and subclonal TP53, SF3B1, BIRC3, NOTCH1, and ATM mutations in chronic lymphocytic leukemia. Blood. 2016;127(17):2122-2130.
27. Nadeu F, Clot G, Delgado J, et al. Clinical impact of the subclonal architecture and mutational complexity in chronic lymphocytic leukemia. Leukemia. 2018;32(3):645-653.
28. Wang J, Khiabanian H, Rossi D, et al. Tumor evolutionary directed graphs and the history of chronic lymphocytic leukemia. Elife. 2014;3:e02869.
29. Jones DT, Hutter B, Jager N, et al. Recurrent somatic alterations of FGFR1 and NTRK2 in pilocytic astrocytoma. Nat Genet. 2013;45(8):927-932.
30. Koboldt DC, Zhang Q, Larson DE, et al. VarScan 2: somatic mutation and copy number alteration discovery in cancer by exome sequencing. Genome Res. 2012;22(3):568-576.
31. Carter SL, Cibulskis K, Helman E, et al. Absolute quantification of somatic DNA alterations in human cancer. Nat Biotechnol. 2012;30(5):413-421.
32. Strino F, Parisi F, Micsinai M, Kluger Y. TrAp: a tree approach for fingerprinting subclonal tumor composition. Nucleic Acids Res. 2013;41(17):e165.
33. Jones DT, Jager N, Kool M, et al. Dissecting the genomic complexity underlying medulloblastoma. Nature. 2012;488(7409):100-105.
34. Bailey MH, Tokheim C, Porta-Pardo E, et al. Comprehensive Characterization of Cancer Driver Genes and Mutations. Cell. 2018;173(2):371-385.

35. Beekman R, Chapaprieta V, Russinol N, et al. The reference epigenome and regulatory chromatin landscape of chronic lymphocytic leukemia. Nat Med. 2018;24(6):868-880.
36. Amin NA, Seymour E, Saiya-Cork K, Parkin B, Shedden K, Malek SN. A Quantitative Analysis of Subclonal and Clonal Gene Mutations before and after Therapy in Chronic Lymphocytic Leukemia. Clin Cancer Res. 2016;22(17):4525-4535.
37. Gruber M, Bozic I, Leshchiner I, et al. Growth dynamics in naturally progressing chronic lymphocytic leukaemia. Nature. 2019;570(7762):474-479.
38. Tausch E, Close W, Dolnik A, et al. Venetoclax resistance and acquired BCL2 mutations in chronic lymphocytic leukemia. Haematologica. 2019;104(9):e434-e437.
39. Blombery P, Anderson MA, Gong JN, et al. Acquisition of the Recurrent Gly101Val Mutation in BCL2 Confers Resistance to Venetoclax in Patients with Progressive Chronic Lymphocytic Leukemia. Cancer Discov. 2019;9(3):342-353.
40. Woyach JA, Furman RR, Liu TM, et al. Resistance mechanisms for the Bruton's tyrosine kinase inhibitor ibrutinib. N Engl J Med. 2014;370(24):2286-2294.

Figure Legends

Figure 1: SNVs and InDels in patients across phases

A) Somatic single nucleotide variants (SNVs) and Insertionen/Deletionen (InDels) identified in CLL samples (SNVs in blue, Indels in green). Genes with recurrent (patients n>2) somatic SNVs and Indels identified in our study or CLL drivers from previously published CLL cohorts highlighted by black boxes (4,26, COSMIC 19/03/14). Only patients with matched control sample were considered.

B) Each symbol shows the number of variants identified in one patient sample. Samples are grouped and colored in dependence of treatment and outcome of the subsequent phase. Black bars represent group means.

Figure 2: Allele frequency changes during the clinical course of CLL patients

Overall clinical phase is depicted by colored bars on top (green=untreated, yellow=relapsed after treatment, red=refractory to treatment). The individual treatments and disease progressions are depicted in the first row of each graph with gray bars representing treatment, while green, yellow, orange and red bars indicate the type of treatment response and disease progression. Richter transformation in HU-1-32 (blue) is highlighted in the second row, above the alternative allele frequencies. Allele frequency changes are colored based on hierarchical clustering of the trajectories following by identifying the six major clusters (R function cutree). Y-axis: indicates allele frequencies, x-axis: indicates time course in years.

Figure 3: Allele frequency changes across phases

Between two consecutive samplings in a patient, changes in allele frequency of clustered somatic SNVs shown as circles. Circles are grouped according to the clinical phase at the second time point. To weigh each patient identically, regardless of the number of mutations, the changes in allele frequency of all SNVs were clustered into six groups per patient and the average of these groups is depicted resulting in six circles per patient phase. Black line represents phase mean. Statistical significance of allele frequency change differences between clinical phases was tested using Kruskal-Wallis test and p-value corrected using 100000 permutations of the phase labels for the six mutation clusters representing each patient ($p= 0.00262$).

Figure 4: Changes in Cancer Cell Fractions and evolution types

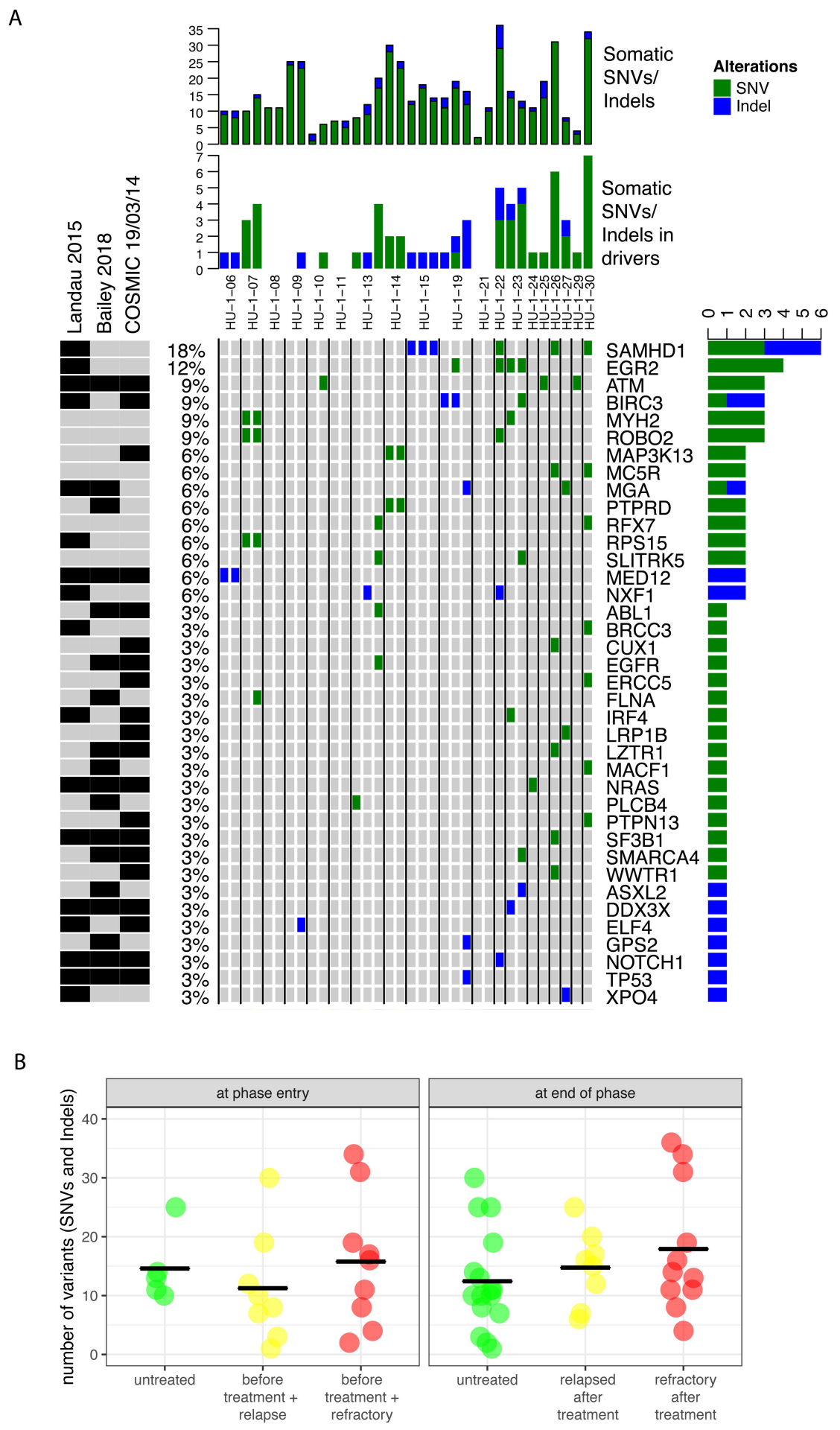
A) Evolution patterns in 12 patients. Probability distribution of the cancer cell fraction (CCF) for each somatic SNV revealed clonal (red) or subclonal (blue) SNVs (left side). The changes in CCF are depicted in the last column. Changes (a mutation with a change in CCF of greater than 0.2 (delta CCF >0.2) with probability >0.5) are highlighted in green (increased CCF) or red (reduced CCF). On the basis of time dependent changes of CCF (right side), evolution patterns were considered as (unbranched) co-evolution (C) when significant changes were unidirectional (up or down), or branched (B) when significant changes were in both directions (up and down) indicating that a dominant clone is replaced by its siblings. Time between samplings is indicated in years at the top. B) Difference in occurrence of evolution types across clinical phases (coevolution = blue, branched evolution = brown).

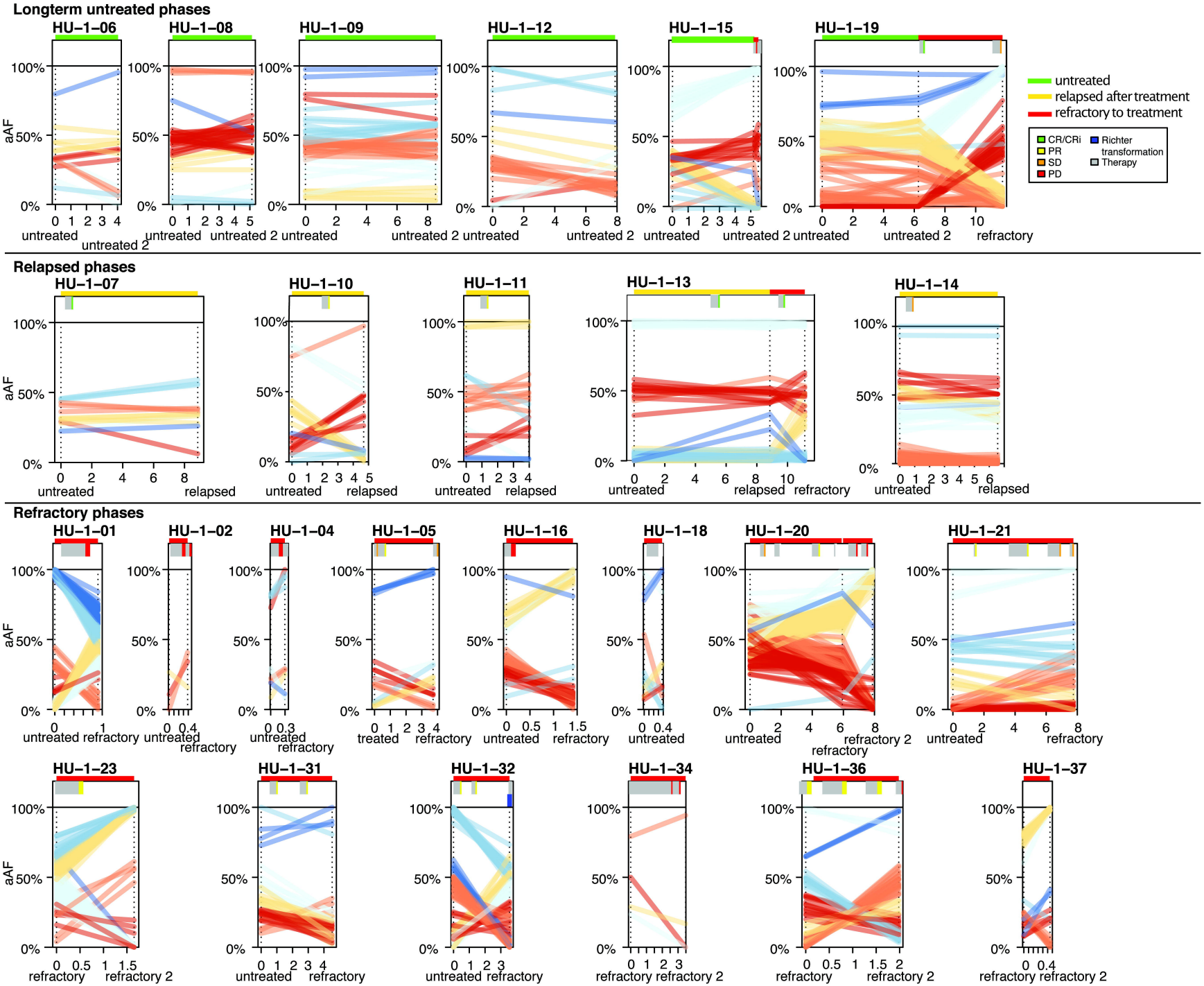
Figure 5: Clonal composition as inferred by TrAP

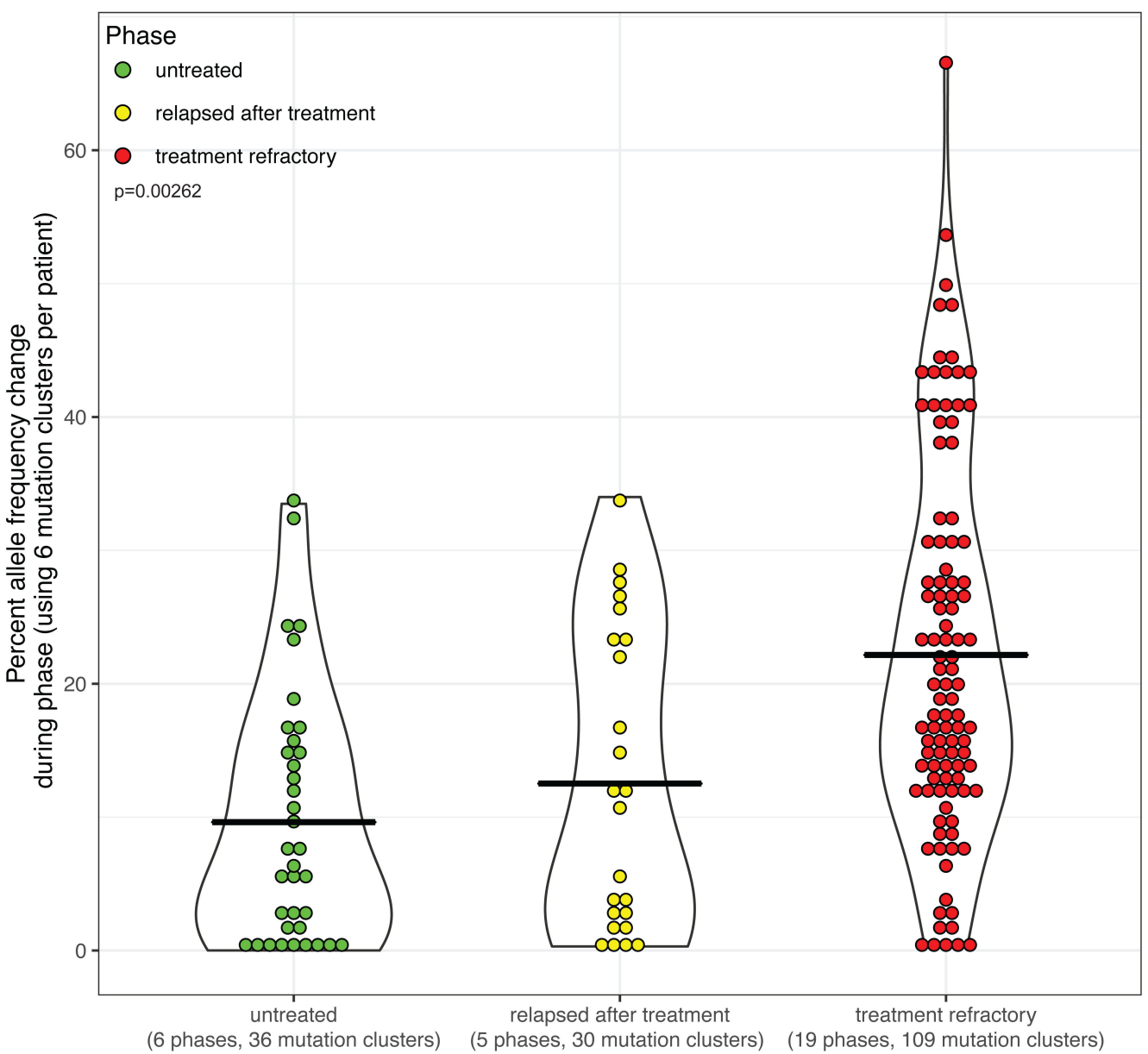
Exemplary changes in clone fraction for individual patients across samplings. Clonal composition was estimated by applying TrAP on the cancer cell fraction calculated for individual SNVs. Different clones are represented by the respective colors and individual time points are indicated on the x axis. Highlighted with the respective gene symbols are inferred clones linked to a known cancer driver gene. Y axis represents clonal fraction of the individual clones identified for the best fit TrAP solution.

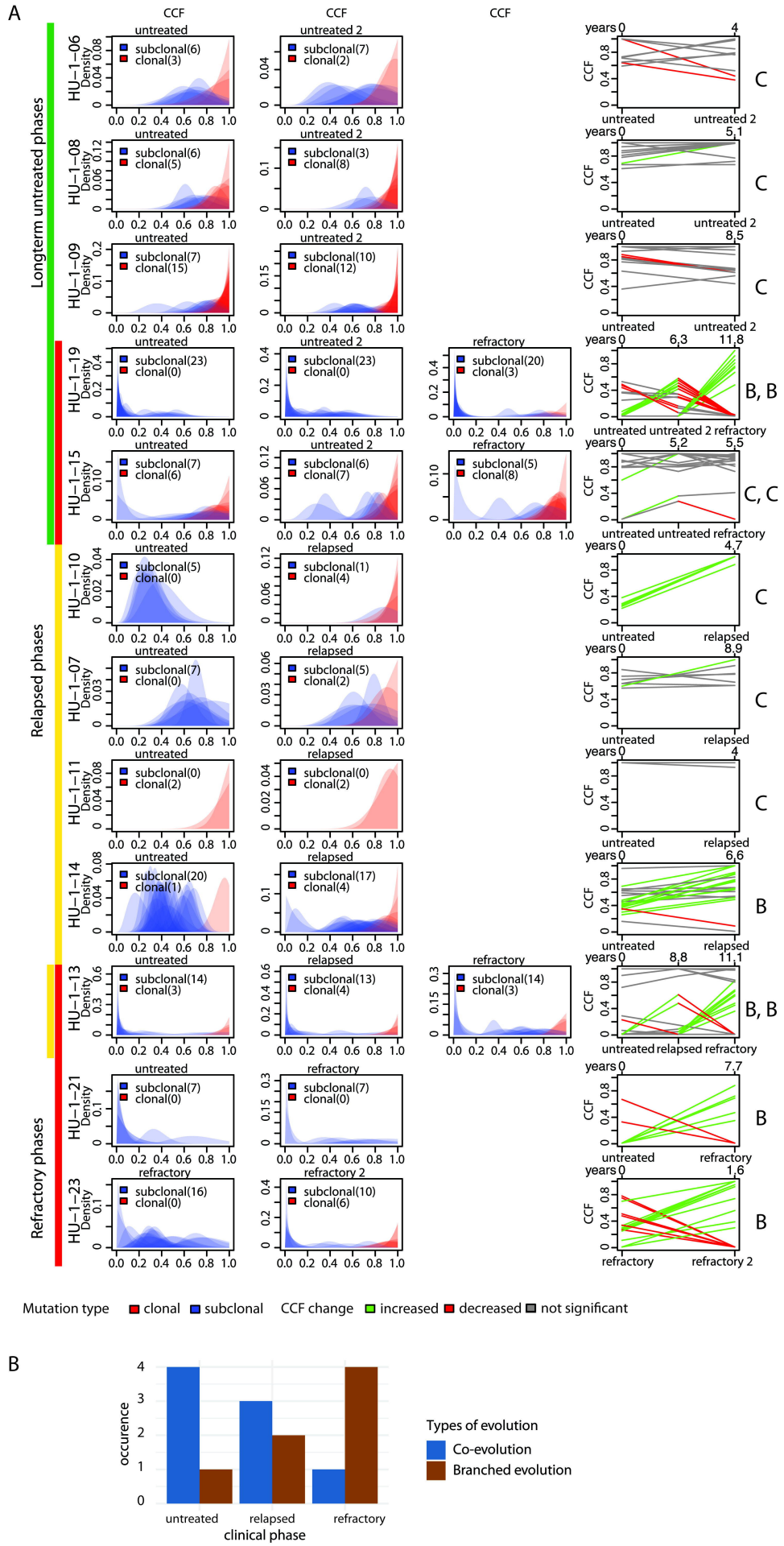
Figure 6: Model of the clonal composition changes

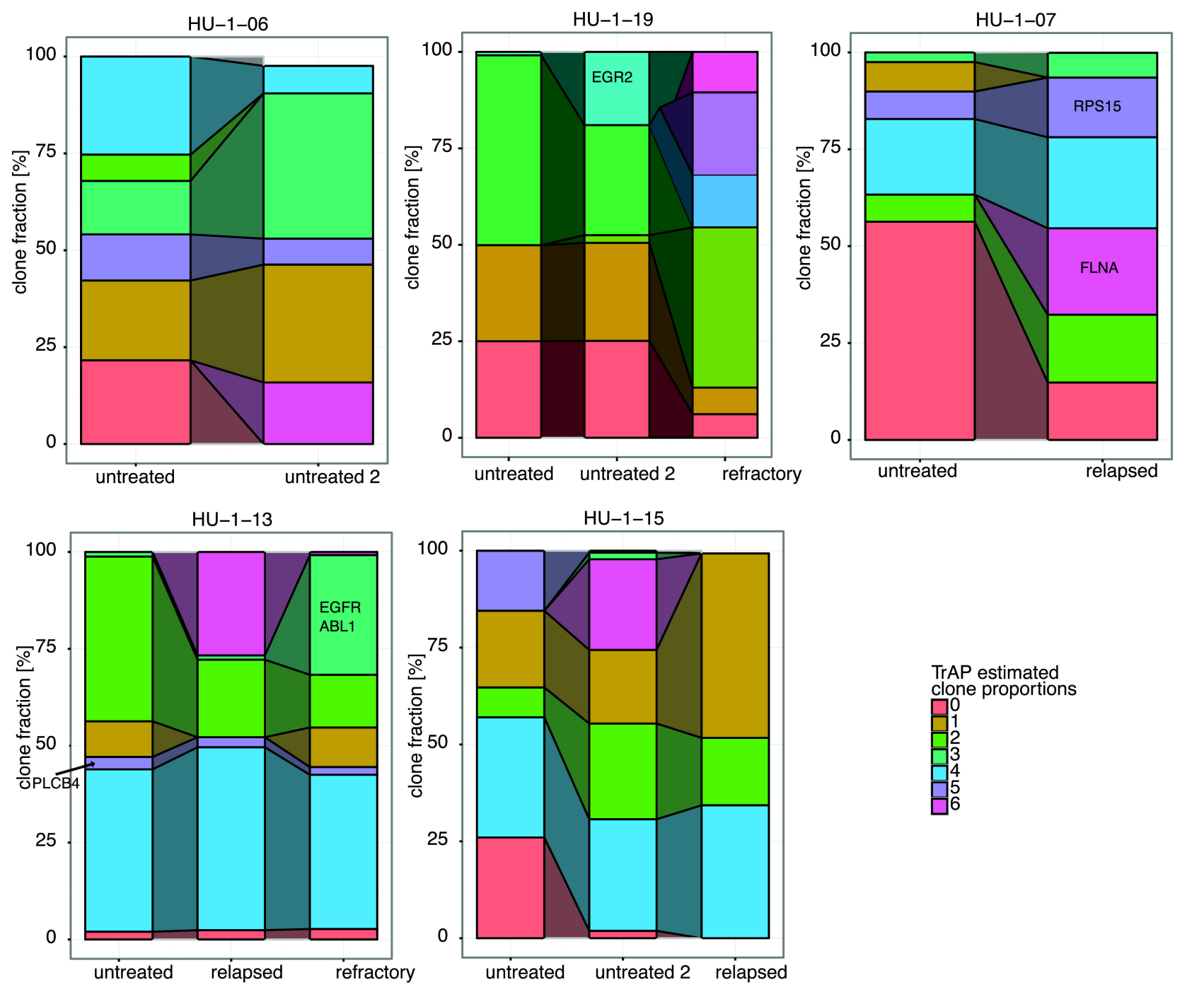
Model of the clonal composition changes in the three different treatment phases (longterm untreated, relapsed and treatment refractory). Black lines indicate lymphocyte counts as surrogate marker for tumor load. Arrows indicate times of treatment. The stacked barplot indicates clonal tumor composition where different colors indicate a different clone defined by a set of mutations.

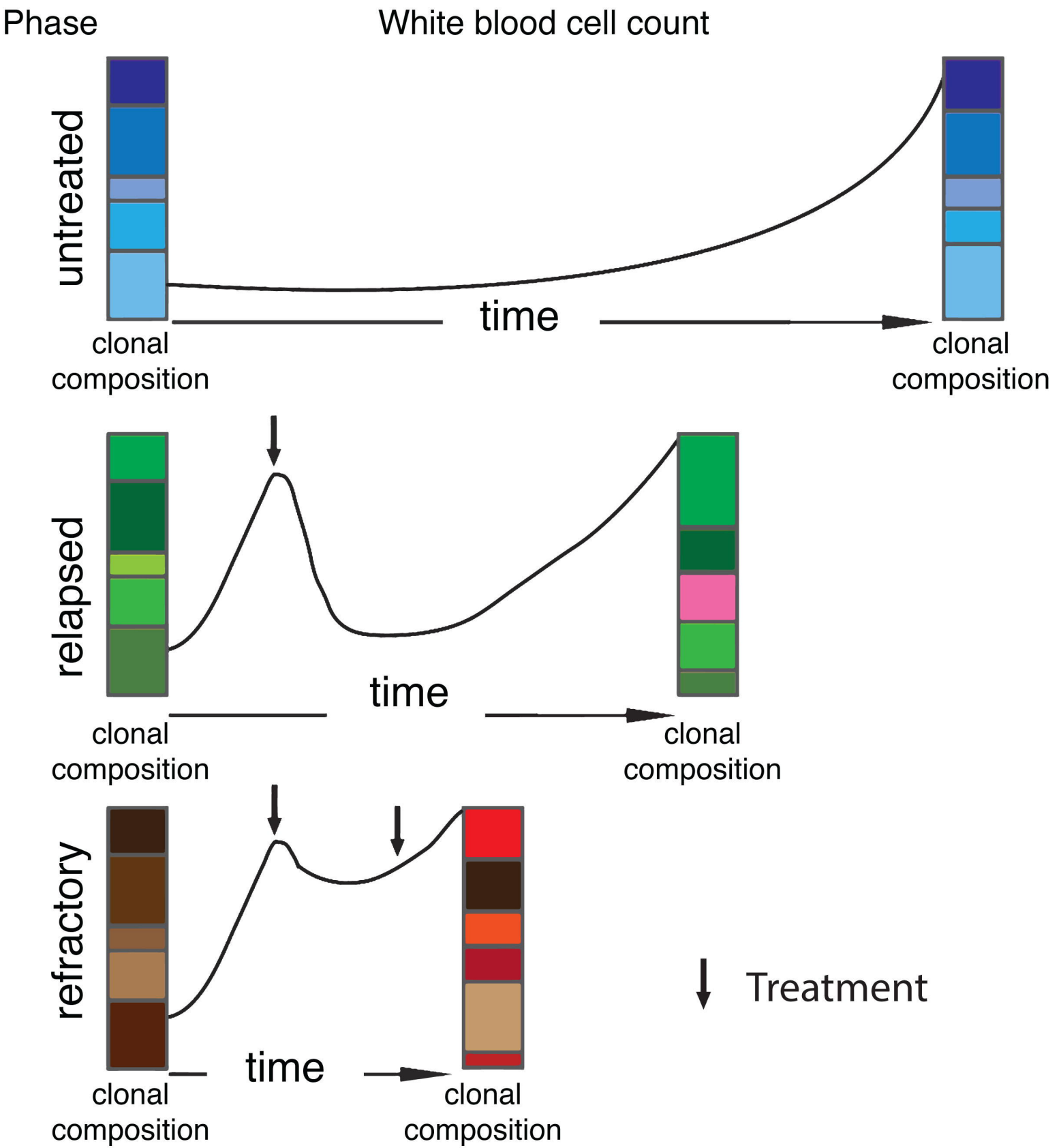

A
Somatic SNVs/ Indels
Somatic SNVs/ Indels in drivers
Alterations
SNV
Indel
Landau 2015
Bailey 2018
COSMIC 19/03/14
HU–1–06
HU–1–07
HU–1–08
HU–1–09
HU–1–10
HU–1–11
HU–1–13
HU–1–14
HU–1–15
HU–1–19
HU–1–21
HU–1–22
HU–1–23
HU–1–24
HU–1–25
HU–1–26
HU–1–27
HU–1–29
HU–1–30
18% SAMHD1
12% EGR2
9% ATM
9% BIRC3
9% MYH2
9% ROBO2
6% MAP3K13
6% MC5R
6% MGA
6% PTPRD
6% RFX7
6% RPS15
6% SLITRK5
6% MED12
6% NXF1
3% ABL1
3% BRCC3
3% CUX1
3% EGFR
3% ERCC5
3% FLNA
3% IRF4
3% LRP1B
3% LZTR1
3% MACF1
3% NRAS
3% PLCB4
3% PTPN13
3% SF3B1
3% SMARCA4
3% WWTR1
3% ASXL2
3% DDX3X
3% ELF4
3% GPS2
3% NOTCH1
3% TP53
3% XPO4
B
at phase entry
at end of phase
number of variants (SNVs and Indels)
untreated
before treatment + relapse
before treatment + refractory
untreated
relapsed after treatment
refractory after treatment


Longterm untreated phases
HU−1−06
HU−1−08
HU−1−09
HU−1−12
HU−1−15
HU−1−19
aAF
100%
50%
0%
untreated
untreated 2
refractory
Relapsed phases
HU−1−07
HU−1−10
HU−1−11
HU−1−13
HU−1−14
relapsed
Refractory phases
HU−1−01
HU−1−02
HU−1−04
HU−1−05
HU−1−16
HU−1−18
HU−1−20
HU−1−21
HU−1−23
HU−1−31
HU−1−32
HU−1−34
HU−1−36
HU−1−37
refractory
refractory 2
untreated
relapsed after treatment
refractory to treatment
CR/CRi
PR
SD
PD
Richter transformation
Therapy


Phase
untreated
relapsed after treatment
treatment refractory
p=0.00262
Percent allele frequency change
during phase (using 6 mutation clusters per patient)
60
40
20
0
untreated
(6 phases, 36 mutation clusters)
relapsed after treatment
(5 phases, 30 mutation clusters)
treatment refractory
(19 phases, 109 mutation clusters)


A
CCF
CCF
CCF
Longterm untreated phases
Relapsed phases
Refractory phases
HU−1−06
HU−1−08
HU−1−09
HU−1−19
HU−1−15
HU−1−10
HU−1−07
HU−1−11
HU−1−14
HU−1−13
HU−1−21
HU−1−23
Density
untreated
untreated 2
relapsed
refractory
refractory 2
subclonal(6)
clonal(3)
subclonal(7)
clonal(2)
subclonal(6)
clonal(5)
subclonal(3)
clonal(8)
subclonal(7)
clonal(15)
subclonal(10)
clonal(12)
subclonal(23)
clonal(0)
subclonal(20)
clonal(3)
subclonal(7)
clonal(6)
subclonal(6)
clonal(7)
subclonal(5)
clonal(8)
subclonal(5)
clonal(0)
subclonal(1)
clonal(4)
subclonal(7)
clonal(0)
subclonal(5)
clonal(2)
subclonal(0)
clonal(2)
subclonal(0)
clonal(2)
subclonal(20)
clonal(1)
subclonal(17)
clonal(4)
subclonal(14)
clonal(3)
subclonal(13)
clonal(4)
subclonal(14)
clonal(3)
subclonal(7)
clonal(0)
subclonal(7)
clonal(0)
subclonal(16)
clonal(0)
subclonal(10)
clonal(6)
CCF
years 0 4
untreated untreated 2
C
years 0 5,1
untreated untreated 2
C
years 0 8,5
untreated untreated 2
C
years 0 6,3 11,8
untreated untreated 2 refractory
B, B
years 0 5,2 5,5
untreated untreated refractory
C, C
years 0 4,7
untreated relapsed
C
years 0 8,9
untreated relapsed
C
years 0 4
untreated relapsed
C
years 0 6,6
untreated relapsed
B
years 0 8,8 11,1
untreated relapsed refractory
B, B
years 0 7,7
untreated refractory
B
years 0 1,6
refractory refractory 2
B
Mutation type
clonal
subclonal
CCF change
increased
decreased
not significant
B
occurence
clinical phase
untreated
relapsed
refractory
Types of evolution
Co-evolution
Branched evolution


HU−1−06
HU−1−19
HU−1−07
HU−1−13
HU−1−15
clone fraction [%]
100
75
50
25
0
untreated
untreated 2
refractory
relapsed
EGR2
RPS15
FLNA
EGFR
ABL1
PLCB4
TrAP estimated
clone proportions
0
1
2
3
4
5
6


Phase
White blood cell count
untreated
time
clonal composition
clonal composition
relapsed
time
clonal composition
clonal composition
refractory
time
clonal composition
clonal composition
Treatment

Supplementary Methods

Alignment and variant calling

Alignment and variant calling were performed as described in [1]. In summary, paired-end DNA sequencing reads were mapped to the hg19 assembly (NCBI build 37.1, downloaded from the UCSC genome browser at http://genome.ucsc.edu/) of the human reference genome using BWA version 0.5.10 [2] and processed with samtools (version 0.1.17) and Picard tools (version 1.61, http://broadinstitute.github.io/picard). Single nucleotide variants (SNVs) and small insertions or deletions (indels) were identified based on our in-house analysis pipeline using samtools mpileup and bcftools. For patients without germline sample we called variants using samtools mpileup and bcftools. SNVs with a genotype quality lower than 80 and with no genotype change between consecutive samples were filtered.

CNV calling and calculation of absolute copy numbers

Estimation of the copy number state based on the exome sequencing data was achieved using Varscan 2 on the target regions [3]. Counting the uniquely mapped reads, log2 ratios were calculated for all segments against the respective germline control. Absolute copy numbers were calculated following[4]. In summary, first the ploidy was identified manually based on the FISH results available (see supplementary figures 3, for depiction of the exome sequencing based log2 ratios in regions analyzed by FISH). Incorporating the clinically assessed sample purity measured using FACS, the log2 ratio derived from the read counts of the exome sequencing were transformed into absolute copy numbers for the respective segments of the individual genomes.

Estimation of clonal composition by TrAP

Changes in clonal tumor composition were calculated integrating the CCF at the respective time points using TrAP [5]. In summary, the individual cancer cell fractions across samples per

patient were hierarchically clustered (R: hclust, complete linkage). The CCF profiles were grouped into six clusters (R: cutree k=6). The resulting mean centroids for the six clusters across timepoints were used as input data for TrAP (datatype GAUSSIAN, the error is estimated based on a normal approximation, see [5] for details). For the cluster the median error was calculated for the TrAP based estimation of the subclonal tumor composition. From the possible solutions the solution with the best fit was selected for each patient.

Quantification of DNA methylation and estimation of correlation between time points

DNA from the first and second time point of 10 patient phases (3 long-term untreated, 2 relapsed and 5 refractory) was bisulfite-converted using the EZ DNA Methylation Kit (Zymo Research) and then whole-genome amplified, fragmented and hybridized to Illumina Infinium HumanMethylation450 BeadChips according to the manufacturer's protocol. After single-base extension and staining, the BeadChips were scanned using an Illumina iScan reader. Quality control of the raw data, preprocessing and basic analysis were performed using R/Bioconductor with the RnBeads package [6]. Illumina probes known to be cross-reactive or overlapping known SNPs [7] were excluded from analysis. This was also done for probes giving unreliable measurements as determined by the Greedycut algorithm implemented in RnBeads. The data from the remaining probes were subjected to background subtraction using the Noob method [8] and beta-mixture quantile normalization (BMIQ [9]). In a subsequent step, probes of non-CpG context, probes binding to sequences on sex chromosomes and probes with low standard deviation were filtered out. The data obtained by the remaining probes were used in downstream analyses. Methylation levels of CpG sites were calculated as β-values (β = intensity of the methylated allele (M)/[intensity of the unmethylated allele (U) + intensity of the methylated allele (M) + 100]. For each patient, the shift in overall methylation was estimated by calculating the square of the Pearson correlation coefficient (R^2) between time points using the 40,000 most variable probes in all available samples. Differential methylation analysis between time points for each of the patient groups was performed using the limma

method as implemented in the RnBeads package. For downstream enrichment analyses, we used only differentially methylated CpGs with absolute difference in mean β-values of at least 0.1 at FDR 0.05. Possible enrichment of hypomethylated and hypermethylated CpGs within particular chromatin states was assessed using the R/Bioconductor package LOLA (Locus Overlap Analysis)[10] and published reference CLL epigenomes1[11]. Before the analysis, the genomic coordinates of the reference chromatin state regions were lifted over from hg38 to hg19 using liftOver. Enrichment analyses at the individual patient level were performed using CpGs with absolute difference in β-values between the two time points of at least 0.3.

1. Jones DT, Hutter B, Jager N, et al. Recurrent somatic alterations of FGFR1 and NTRK2 in pilocytic astrocytoma. Nat Genet. 2013;45(8):927-932.
2. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics. 2009;25(14):1754-1760.
3. Koboldt DC, Zhang Q, Larson DE, et al. VarScan 2: somatic mutation and copy number alteration discovery in cancer by exome sequencing. Genome Res. 2012;22(3):568-576.
4. Carter SL, Cibulskis K, Helman E, et al. Absolute quantification of somatic DNA alterations in human cancer. Nat Biotechnol. 2012;30(5):413-421.
5. Strino F, Parisi F, Micsinai M, Kluger Y. TrAp: a tree approach for fingerprinting subclonal tumor composition. Nucleic Acids Res. 2013;41(17):e165.
6. Assenov Y, Muller F, Lutsik P, Walter J, Lengauer T, Bock C. Comprehensive analysis of DNA methylation data with RnBeads. Nat Methods. 2014;11(11):1138-1140.
7. Price ME, Cotton AM, Lam LL, et al. Additional annotation enhances potential for biologically-relevant analysis of the Illumina Infinium HumanMethylation450 BeadChip array. Epigenetics Chromatin. 2013;6(1):4.
8. Triche TJ, Jr., Weisenberger DJ, Van Den Berg D, Laird PW, Siegmund KD. Low-level processing of Illumina Infinium DNA Methylation BeadArrays. Nucleic Acids Res. 2013;41(7):e90.
9. Teschendorff AE, Marabita F, Lechner M, et al. A beta-mixture quantile normalization method for correcting probe design bias in Illumina Infinium 450 k DNA methylation data. Bioinformatics. 2013;29(2):189-196.
10. Sheffield NC, Bock C. LOLA: enrichment analysis for genomic region sets and regulatory elements in R and Bioconductor. Bioinformatics. 2016;32(4):587-589.
11. Beekman R, Chapaprieta V, Russinol N, et al. The reference epigenome and regulatory chromatin landscape of chronic lymphocytic leukemia. Nat Med. 2018;24(6):868-880.

Supplementary Information

Supplementary tables

Suppl. table 1:

Clinical information for the patient samples

Suppl. table 2:

Treatment and sampling information

Suppl. table 3:

Table of driver events

Suppl. table 4:

Evolution types across clinical phases

Suppl. table 5:

Table of SNVs identified

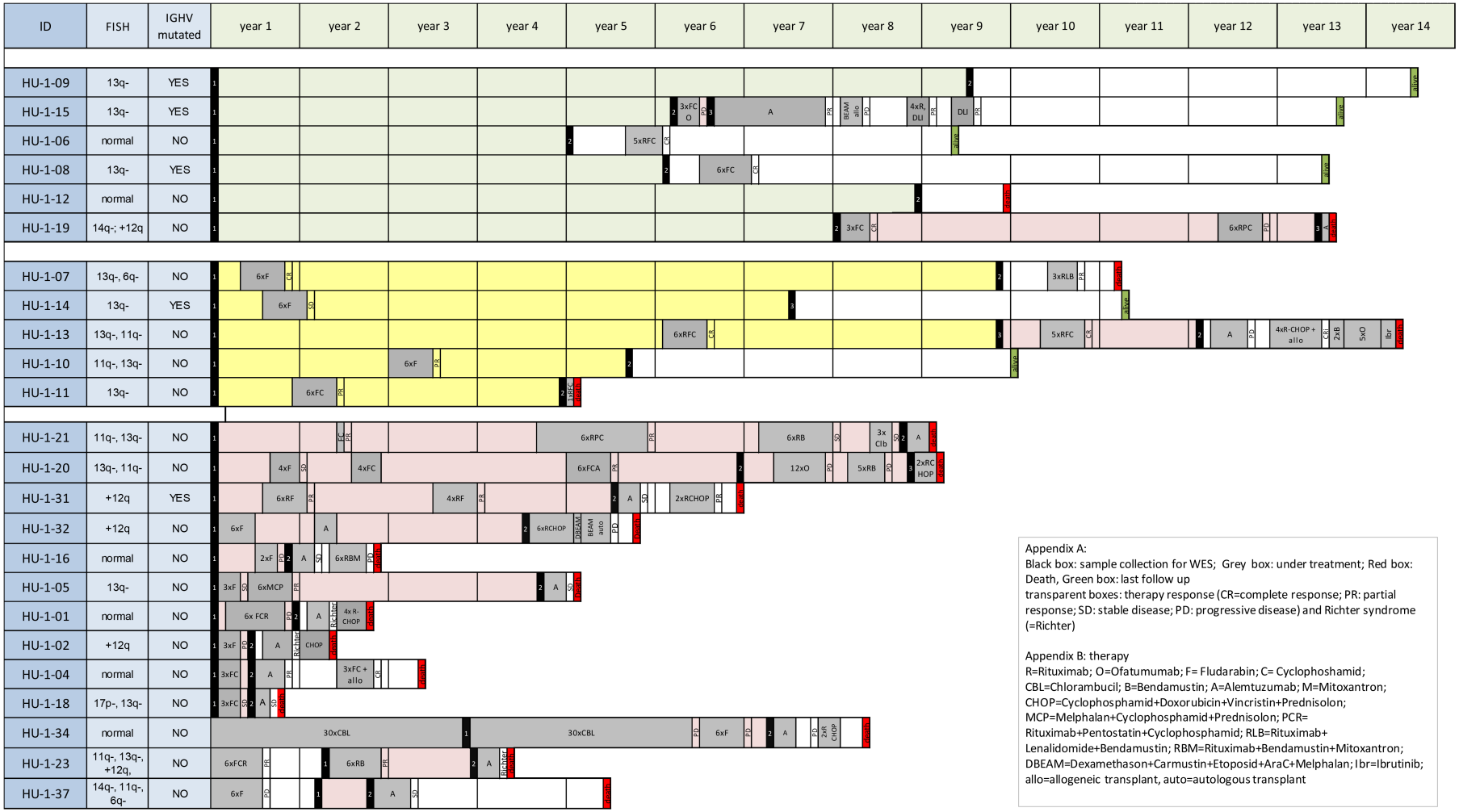
Suppl. table 6:

Representation of samples and figures explaining which samples were used in the different analyses

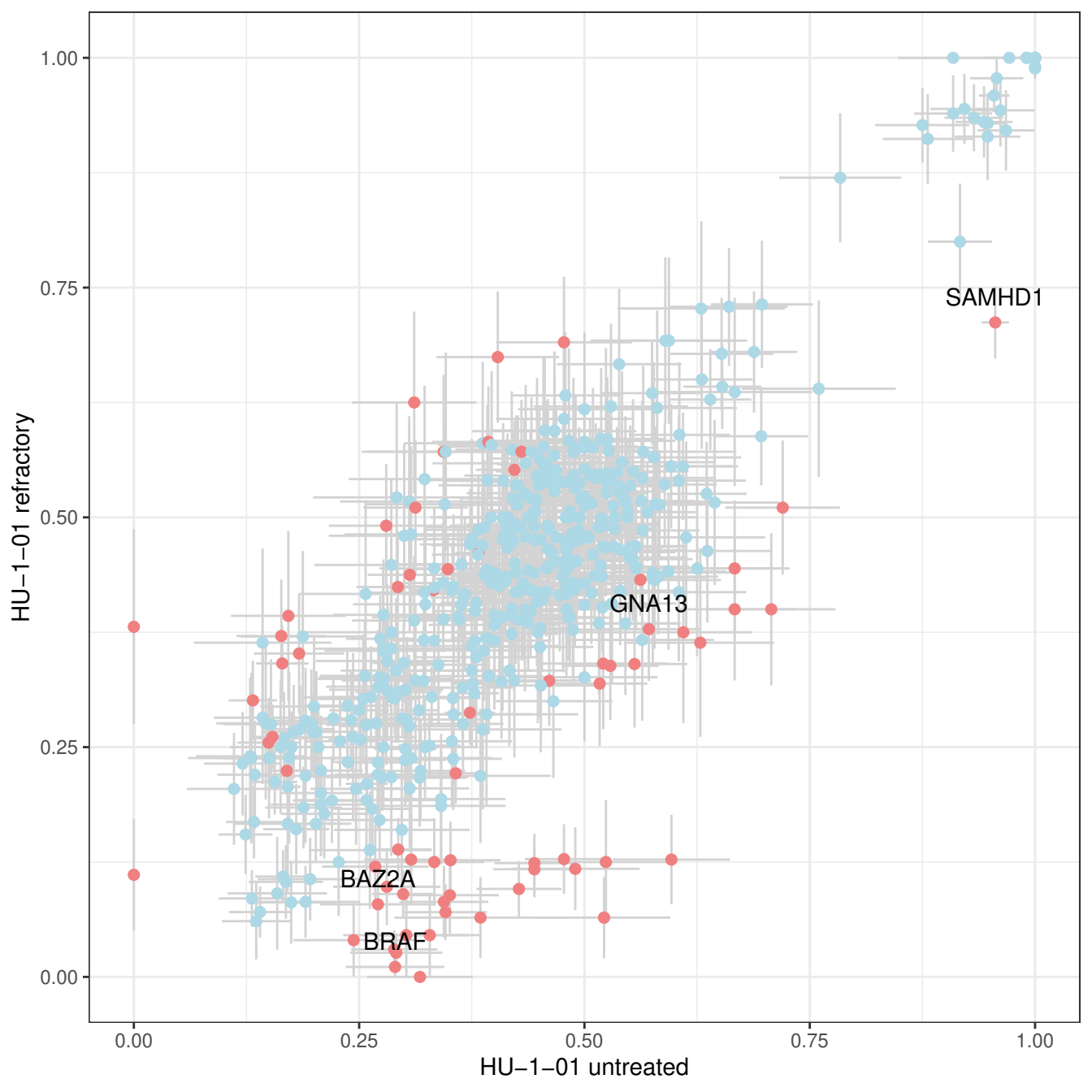
Supplementary figures

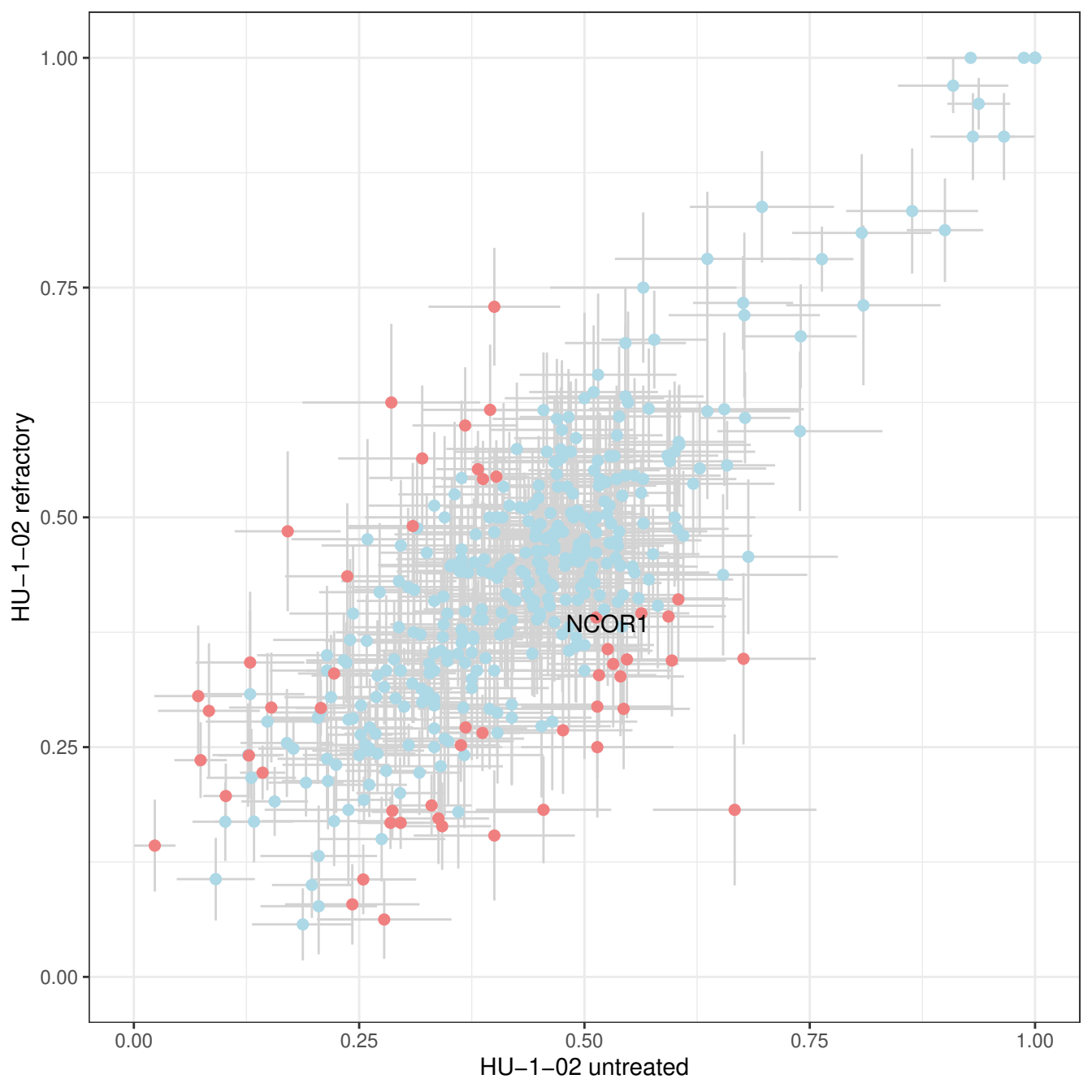
Suppl. figure 1: Graphical representation of clinical course of analyzed patients. Three patients are lost to follow-up (HU-1-09, -13, -15). (Black box: sample collection for WES; Grey box: under treatment; Red box: Death, Green box: last follow up;

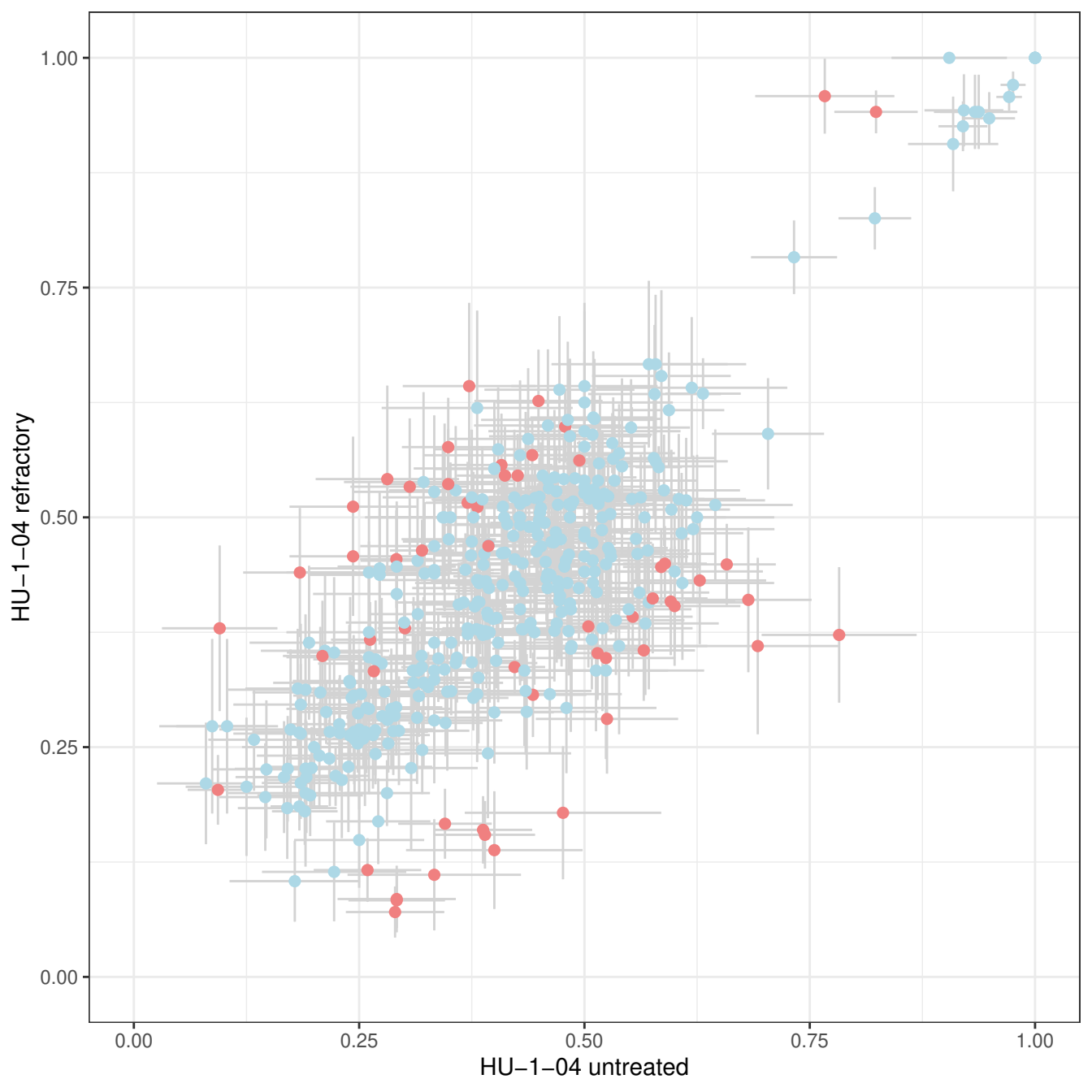
Transparent boxes: therapy response (CR = complete response; PR: partial response; SD: stable disease; PD: progressive disease) and Richter syndrome (= Richter); Therapy information: R = Rituximab; O = Ofatumumab; F = Fludarabin; C = Cyclophoshamid; CBL=Chlorambucil; B=Bendamustin; A=Alemtuzumab; M=Mitoxantron; CHOP = Cyclophosphamid + Doxorubicin + Vincristin + Prednisolon; MCP = Melphalan + Cyclophosphamid + Prednisolon; PCR = Rituximab + Pentostatin + Cyclophosphamid; RLB = Rituximab + Lenalidomide + Bendamustin; RBM= Rituximab + Bendamustin + Mitoxantron; DBEAM = Dexamethason + Carmustin + Etoposid + AraC + Melphalan; Ibr = Ibrutinib; allo = allogeneic transplant, auto = autologous transplant, DLI = Donor lymphocyte infusion)

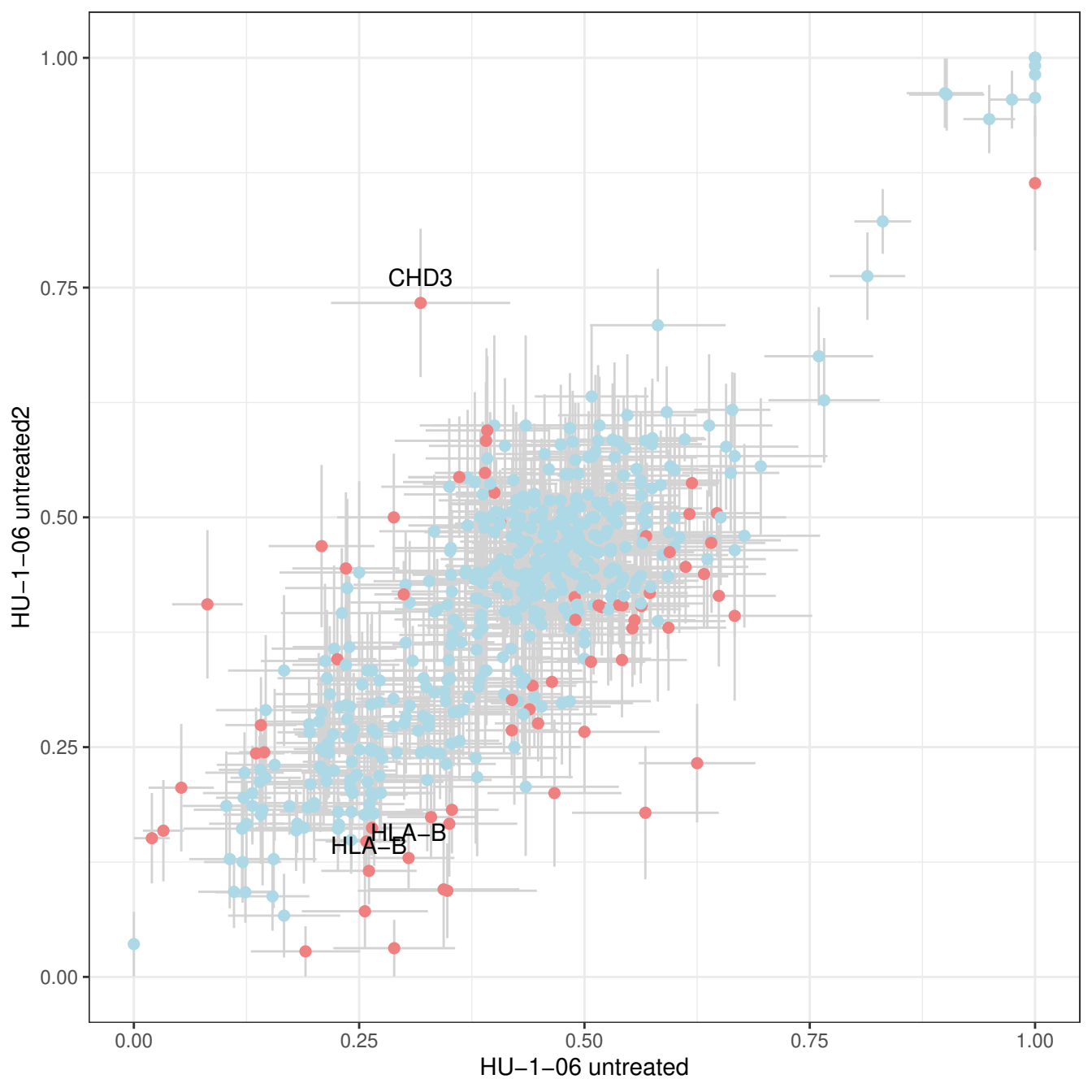

ID
FISH
IGHV mutated
year 1
year 2
year 3
year 4
year 5
year 6
year 7
year 8
year 9
year 10
year 11
year 12
year 13
year 14
HU-1-09 13q- YES
HU-1-15 13q- YES
HU-1-06 normal NO
HU-1-08 13q- YES
HU-1-12 normal NO
HU-1-19 14q-; +12q NO
HU-1-07 13q-, 6q- NO
HU-1-14 13q- YES
HU-1-13 13q-, 11q- NO
HU-1-10 11q-, 13q- NO
HU-1-11 13q- NO
HU-1-21 11q-, 13q- NO
HU-1-20 13q-, 11q- NO
HU-1-31 +12q YES
HU-1-32 +12q NO
HU-1-16 normal NO
HU-1-05 13q- NO
HU-1-01 normal NO
HU-1-02 +12q NO
HU-1-04 normal NO
HU-1-18 17p-, 13q- NO
HU-1-34 normal NO
HU-1-23 11q-, 13q-, +12q, NO
HU-1-37 14q-, 11q-, 6q- NO
6xF
3xFC O
A
4xR, DLI
DLI
5xRFC
6xFC
3xFC
6xRPC
3xRLB
6xF
6xRFC
5xRFC
4xR-CHOP + allo
2xB
5xO
Ibr
6xF
6xFC
6xRPC
6xRB
3x Clb
4xF
4xFC
6xFCA
12xO
5xRB
2xRCHOP
6xRF
4xRF
2xRCHOP
6xF
6xRCHOP
DBEAM
BEAM
allo
2xF
6xRBM
3xF
6xMCP
6x FCR
4x R-CHOP
3xF
CHOP
3xFC + allo
3xFC
30xCBL
30xCBL
6xF
2xR CHOP
6xFCR
6xRB
6xF
alive
death
Appendix A:
Black box: sample collection for WES; Grey box: under treatment; Red box: Death, Green box: last follow up
transparent boxes: therapy response (CR=complete response; PR: partial response; SD: stable disease; PD: progressive disease) and Richter syndrome (=Richter)
Appendix B: therapy
R=Rituximab; O=Ofatumumab; F= Fludarabin; C= Cyclophoshamid; CBL=Chlorambucil; B=Bendamustin; A=Alemtuzumab; M=Mitoxantron; CHOP=Cyclophosphamid+Doxorubicin+Vincristin+Prednisolon; MCP=Melphalan+Cyclophosphamid+Prednisolon; PCR= Rituximab+Pentostatin+Cyclophosphamid; RLB=Rituximab+ Lenalidomide+Bendamustin; RBM=Rituximab+Bendamustin+Mitoxantron; DBEAM=Dexamethason+Carmustin+Etoposid+AraC+Melphalan; Ibr=Ibrutinib; allo=allogeneic transplant, auto=autologous transplant

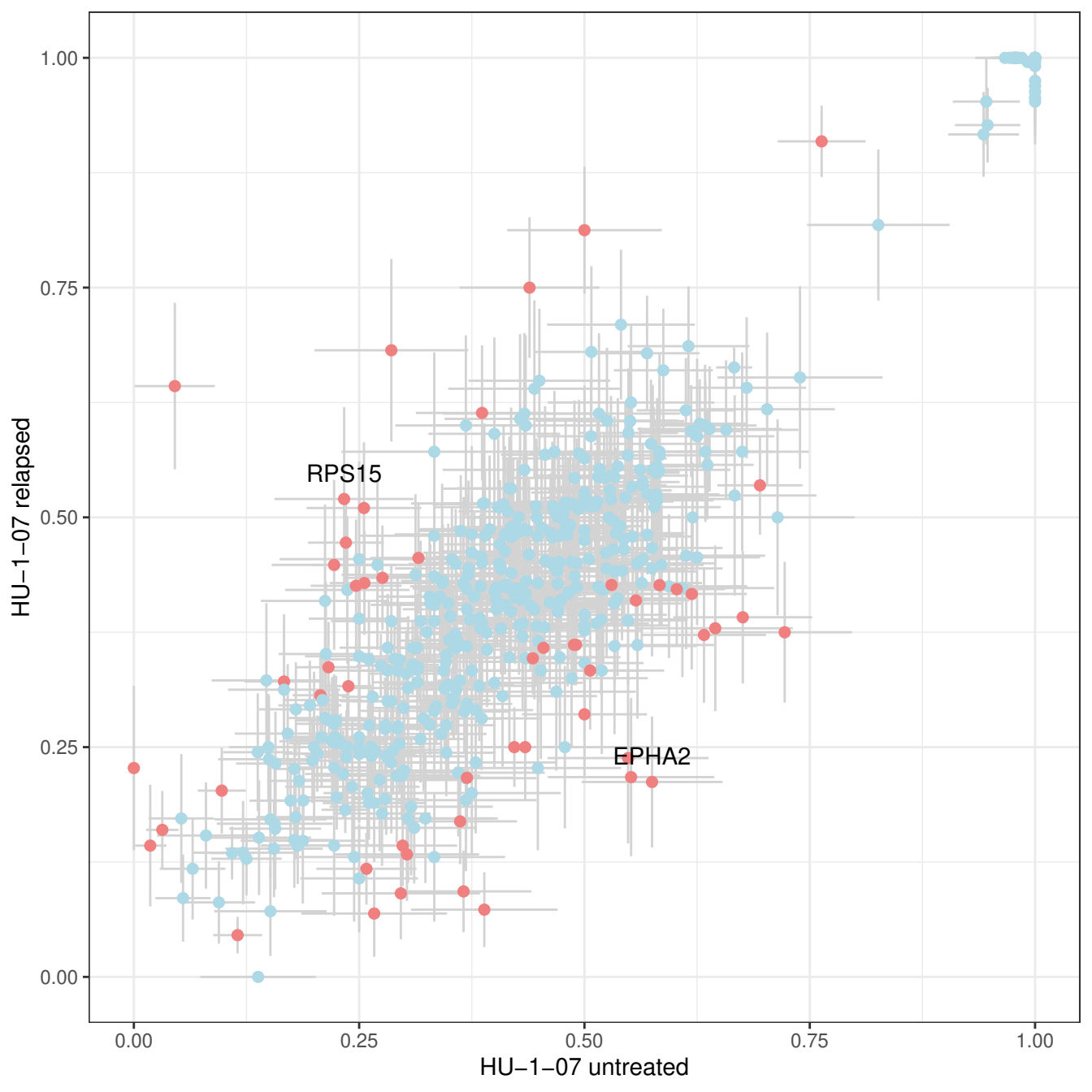
Suppl. figure 2: Scatter plot of variant allele frequency changes between consecutive time points. Known cancer and CLL driver genes (Landau et al. 2015, Bailey et al. 2018) are highlighted in lightred if their change in allele frequency is significant based on Fisher's test on the reference and alternative allele counts other SNVs presented in blue. Excluded are SNVs which are in dbSNP, have a coverage below 20 or a higher frequency than 5 % in 1000 genomes. Gray lines represent error bars for the allele frequency calculation estimated from the read counts.

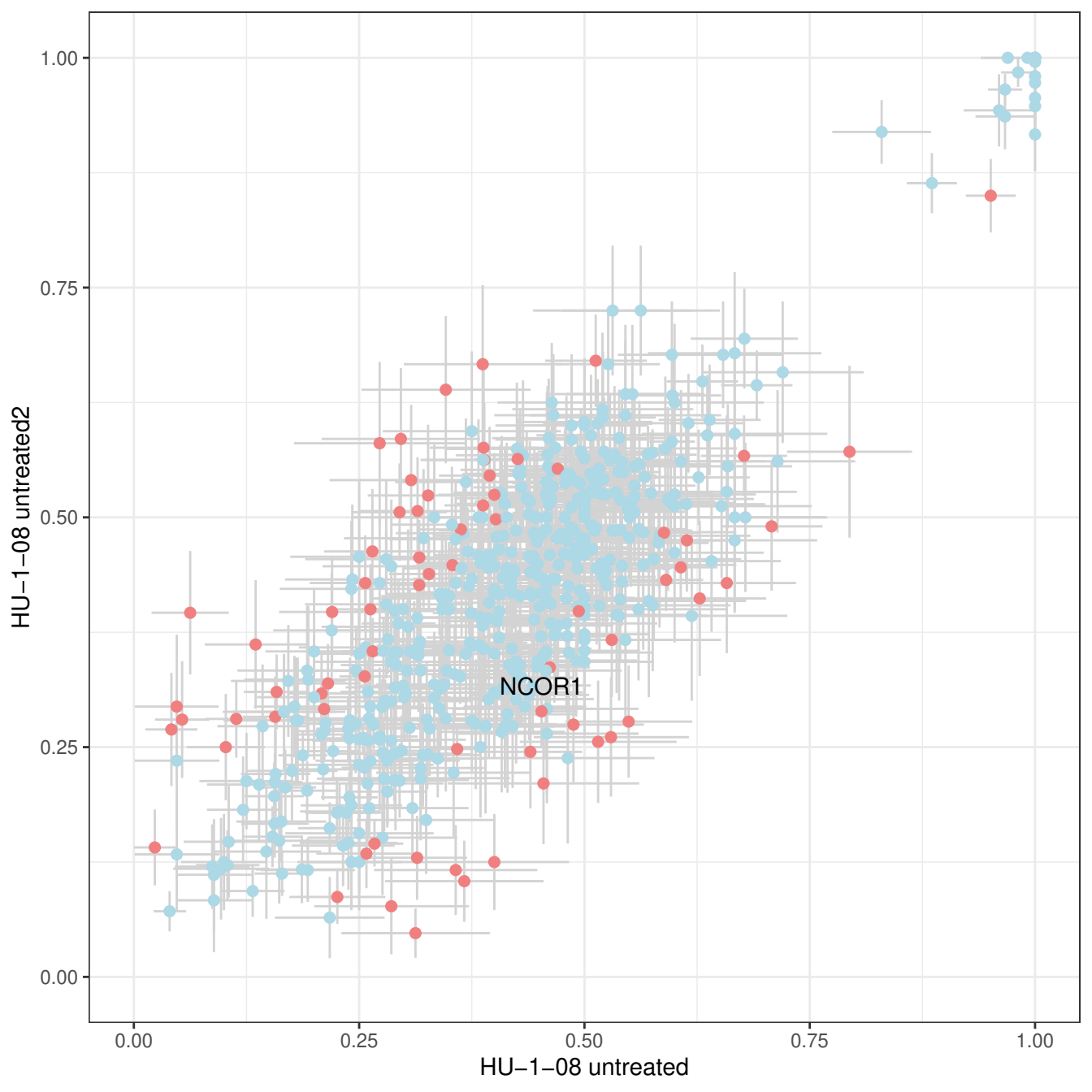

1.00
0.75
0.50
0.25
0.00
HU−1−01 refractory
0.00
0.25
0.50
0.75
1.00
HU−1−01 untreated
SAMHD1
GNA13
BAZ2A
BRAF

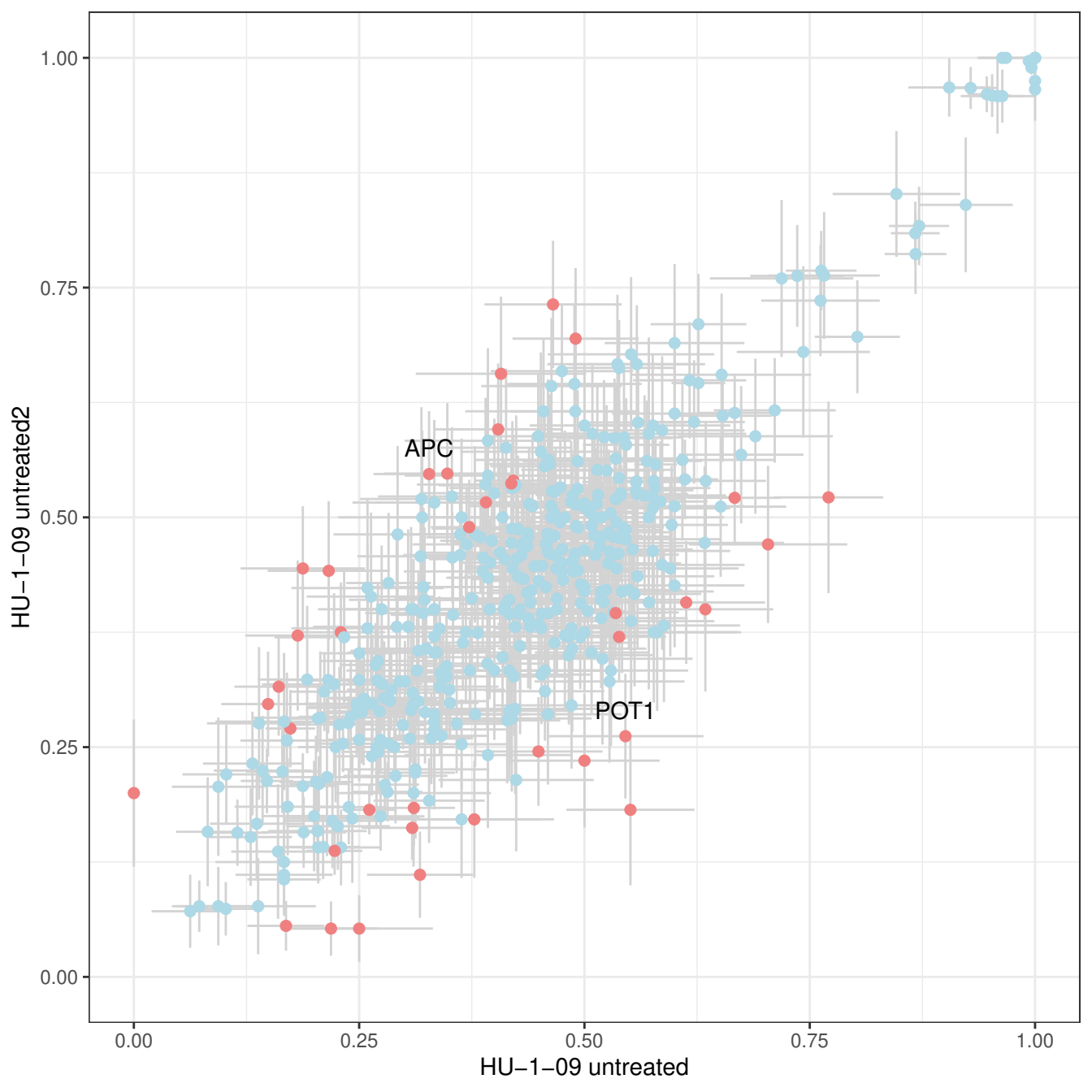

1.00
0.75
0.50
0.25
0.00
HU−1−02 refractory
0.00
0.25
0.50
0.75
1.00
HU−1−02 untreated
NCOR1

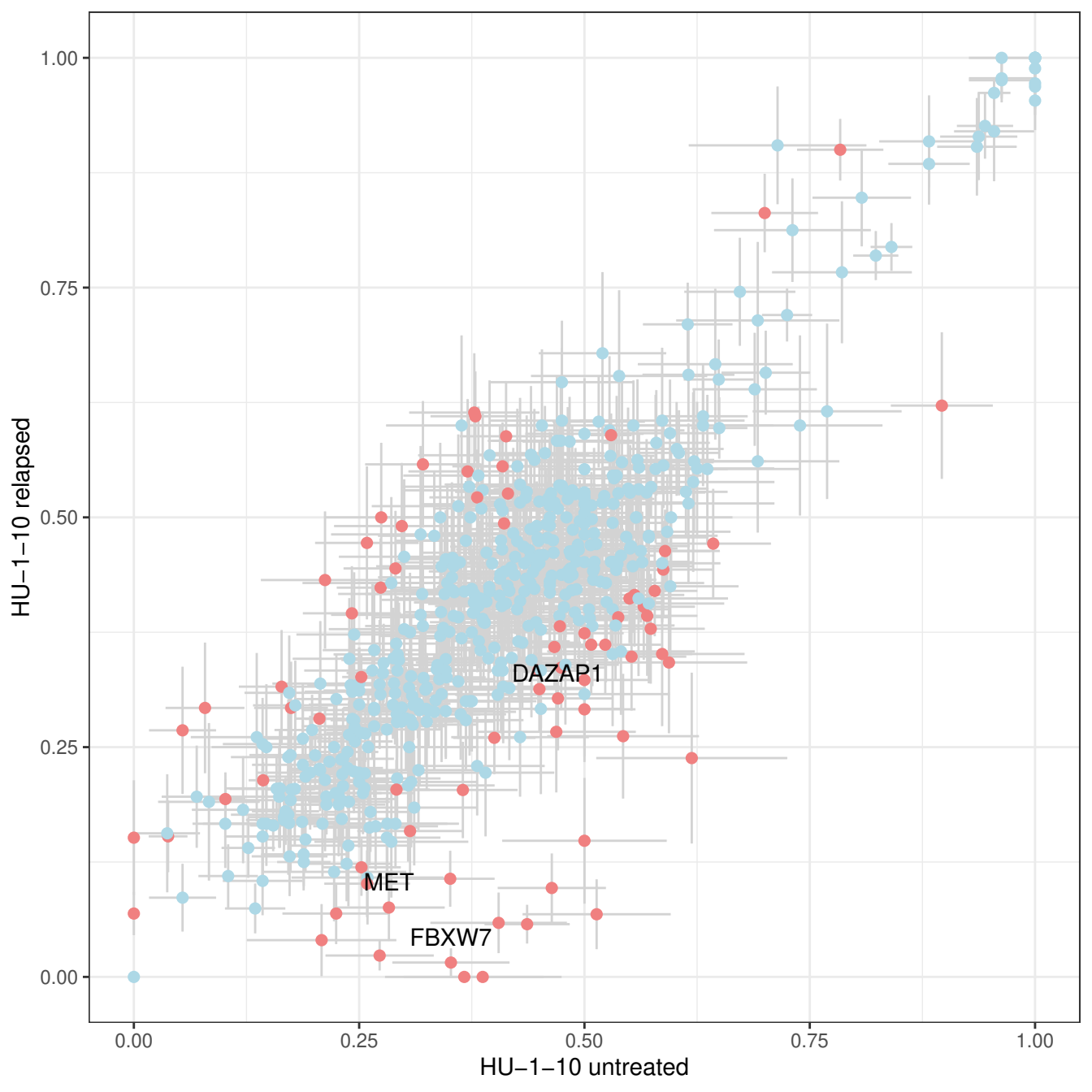

1.00
0.75
0.50
0.25
0.00
HU−1−04 refractory
0.00
0.25
0.50
0.75
1.00
HU−1−04 untreated

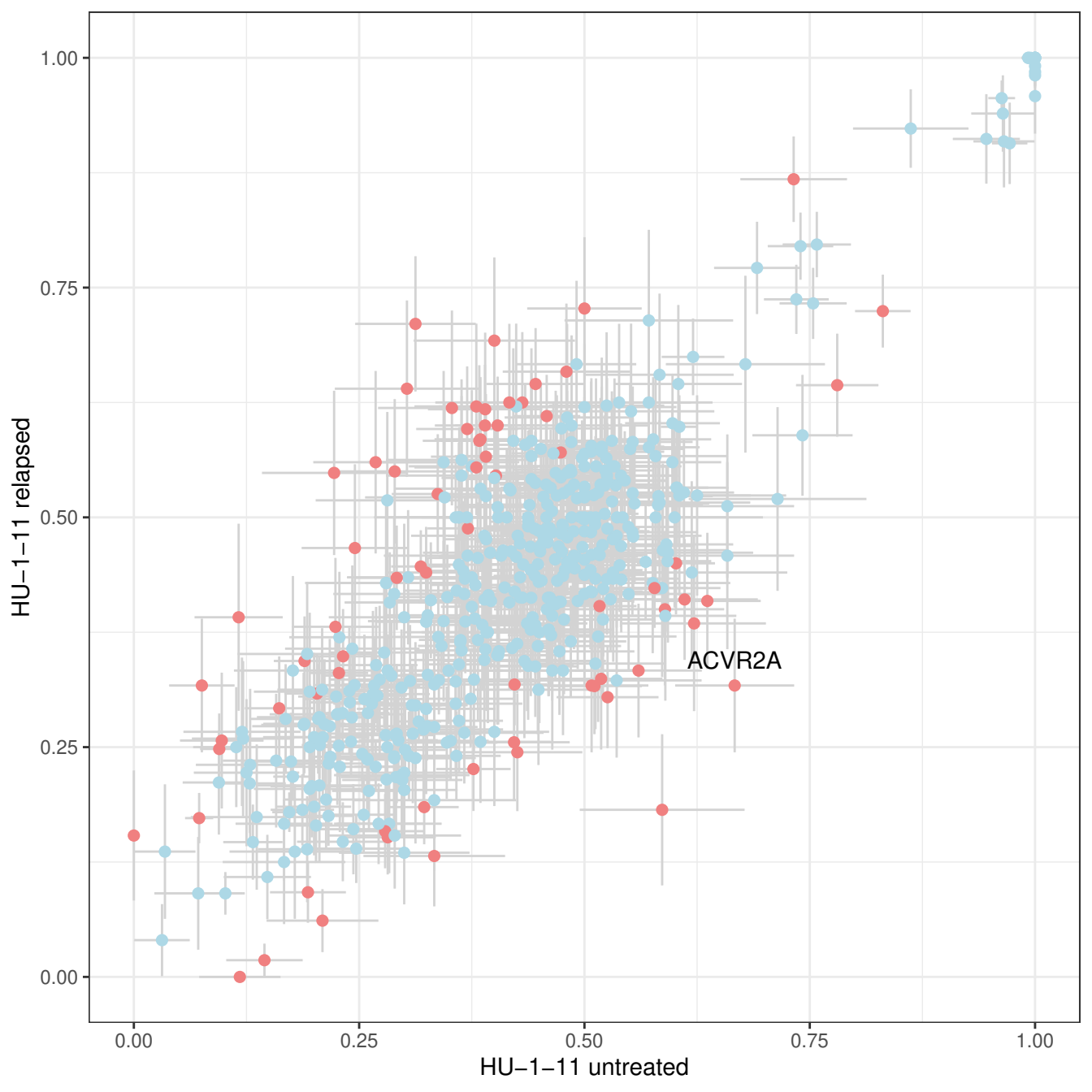

CHD3
HLA−B
HLA−B
1.00
0.75
0.50
0.25
0.00
0.00
0.25
0.50
0.75
1.00
HU−1−06 untreated2
HU−1−06 untreated

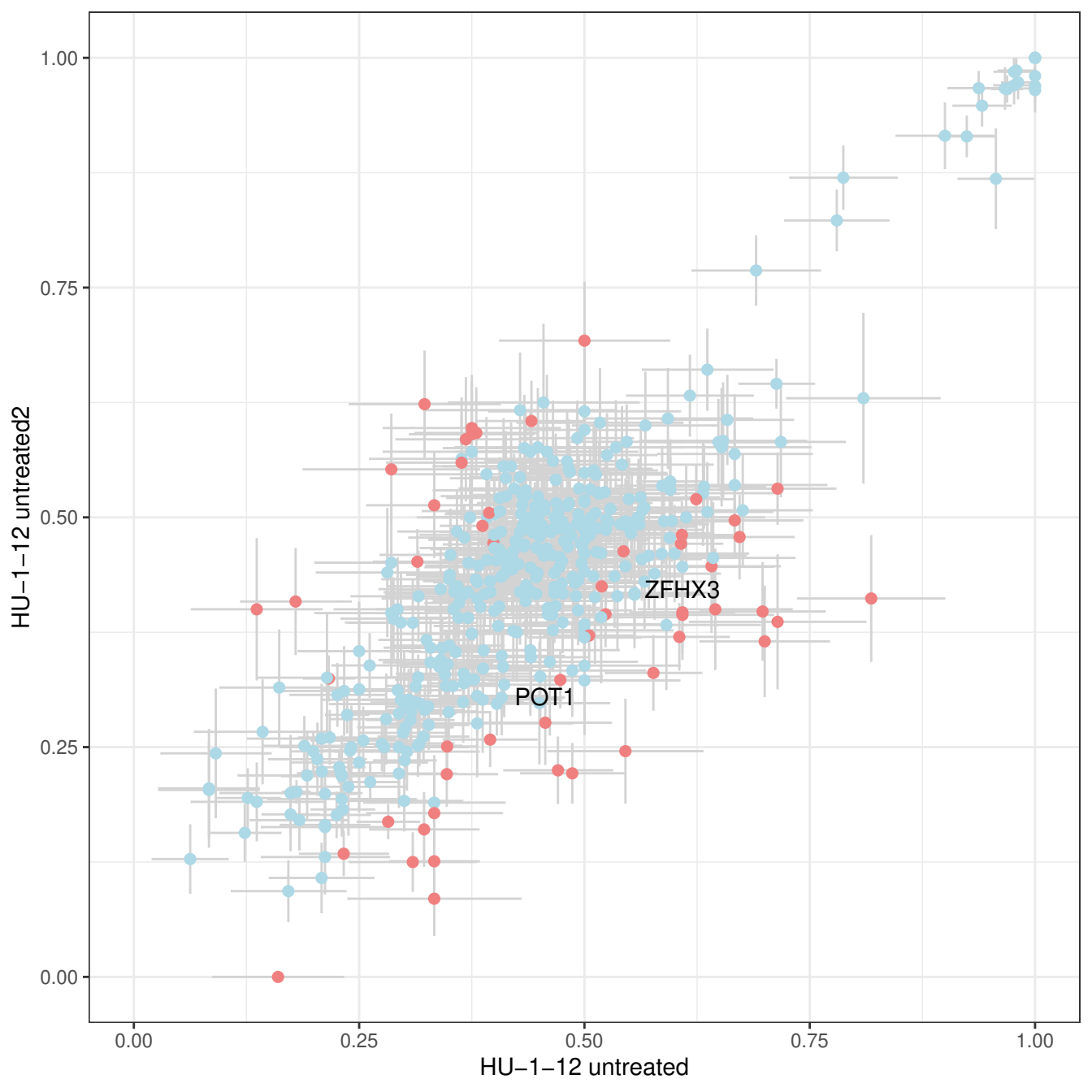

1.00
0.75
0.50
0.25
0.00
HU−1−07 relapsed
RPS15
EPHA2
0.00
0.25
0.50
0.75
1.00
HU−1−07 untreated

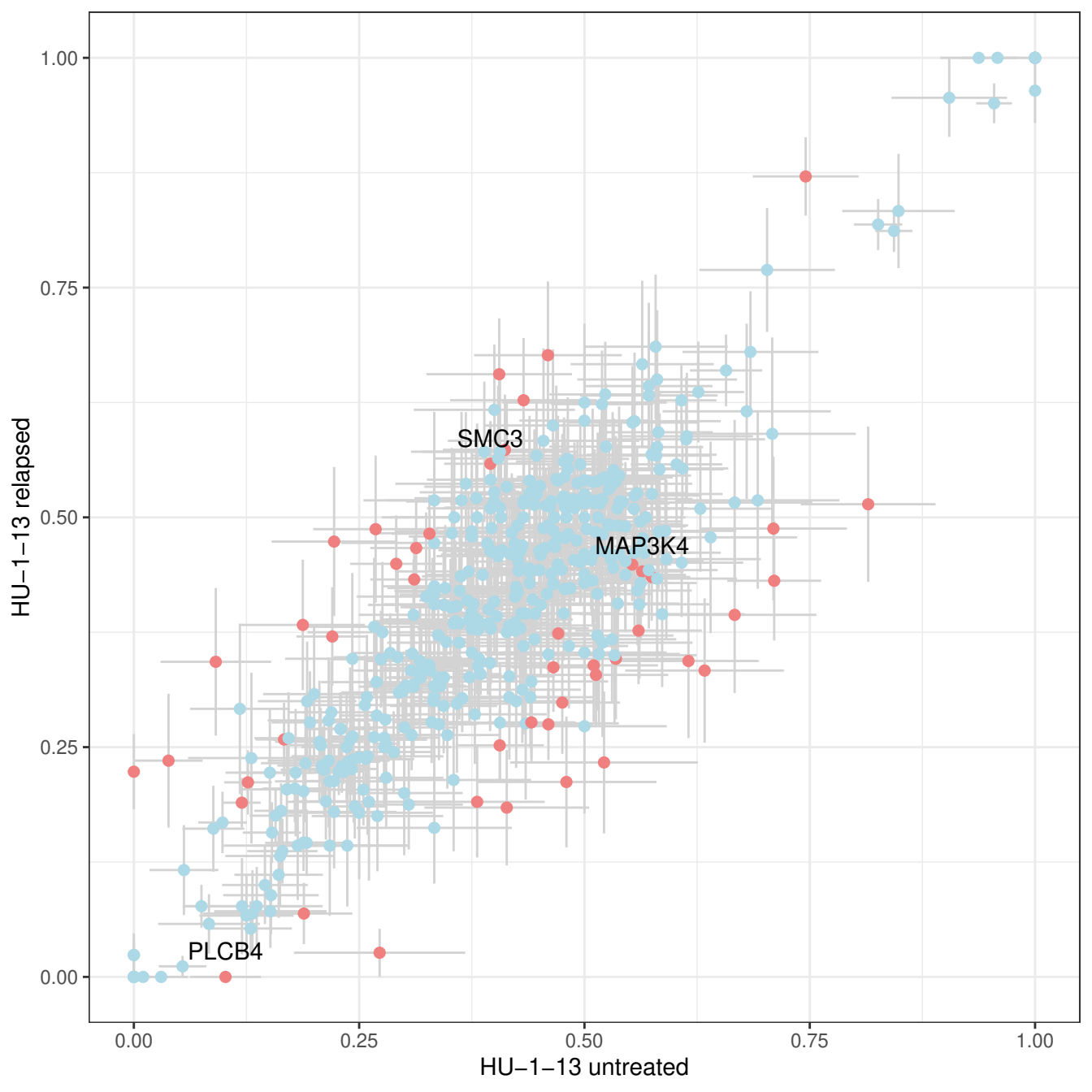

1.00
0.75
0.50
0.25
0.00
HU−1−08 untreated2
0.00
0.25
0.50
0.75
1.00
HU−1−08 untreated
NCOR1

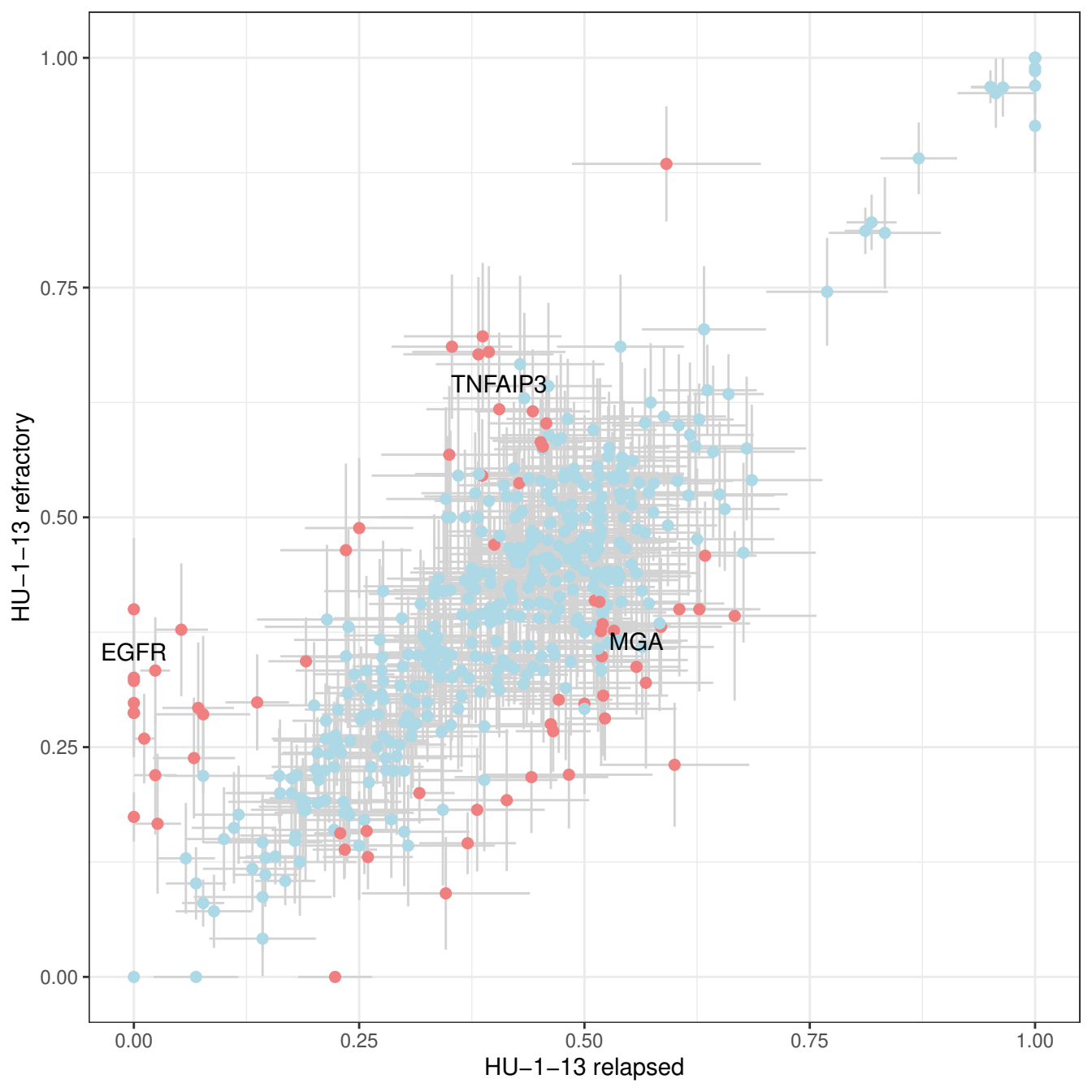

APC
POT1
HU−1−09 untreated2
HU−1−09 untreated
1.00
0.75
0.50
0.25
0.00
0.00
0.25
0.50
0.75
1.00

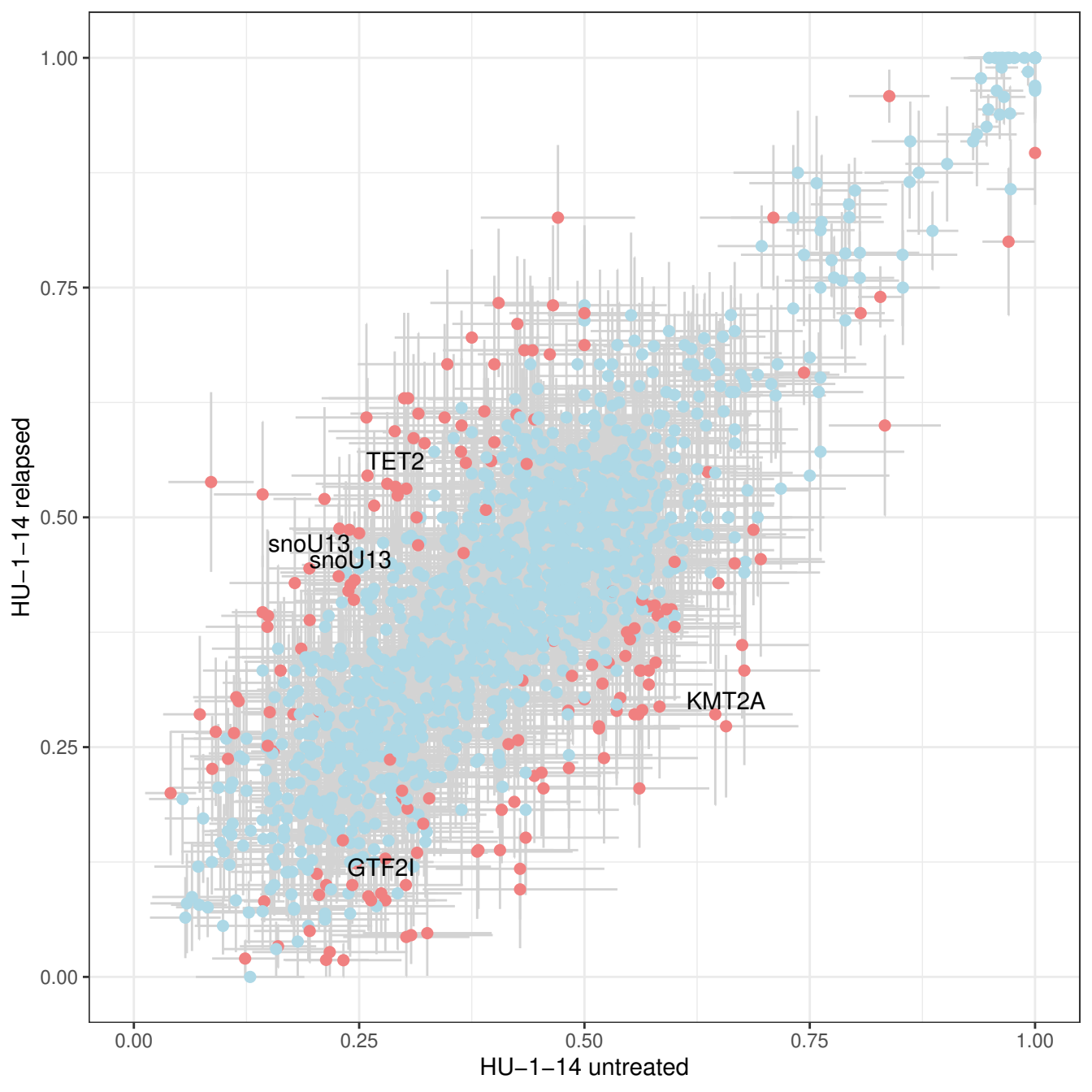

1.00
0.75
0.50
0.25
0.00
HU−1−10 relapsed
0.00
0.25
0.50
0.75
1.00
HU−1−10 untreated
DAZAP1
MET
FBXW7

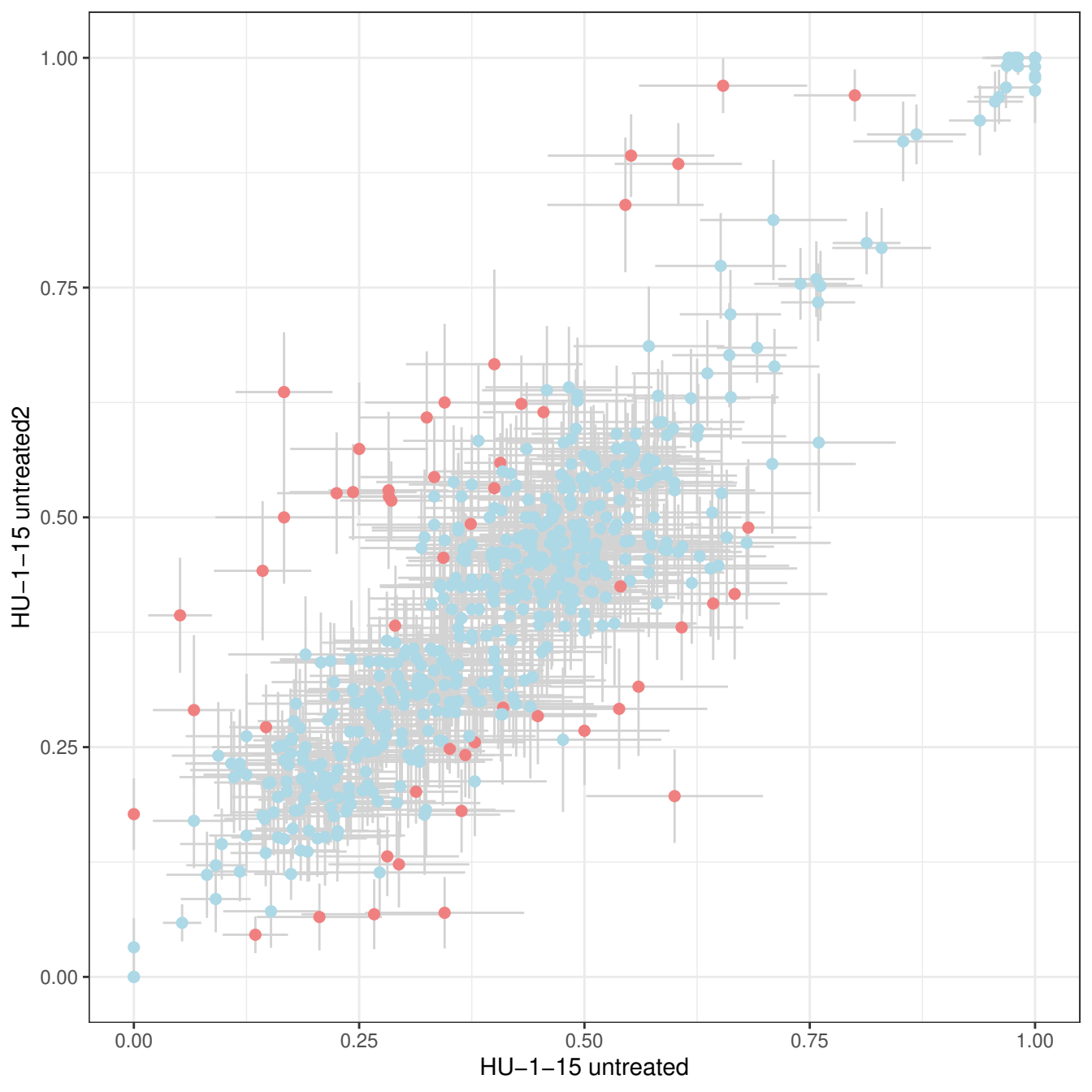

1.00
0.75
0.50
0.25
0.00
HU−1−11 relapsed
0.00
0.25
0.50
0.75
1.00
HU−1−11 untreated
ACVR2A

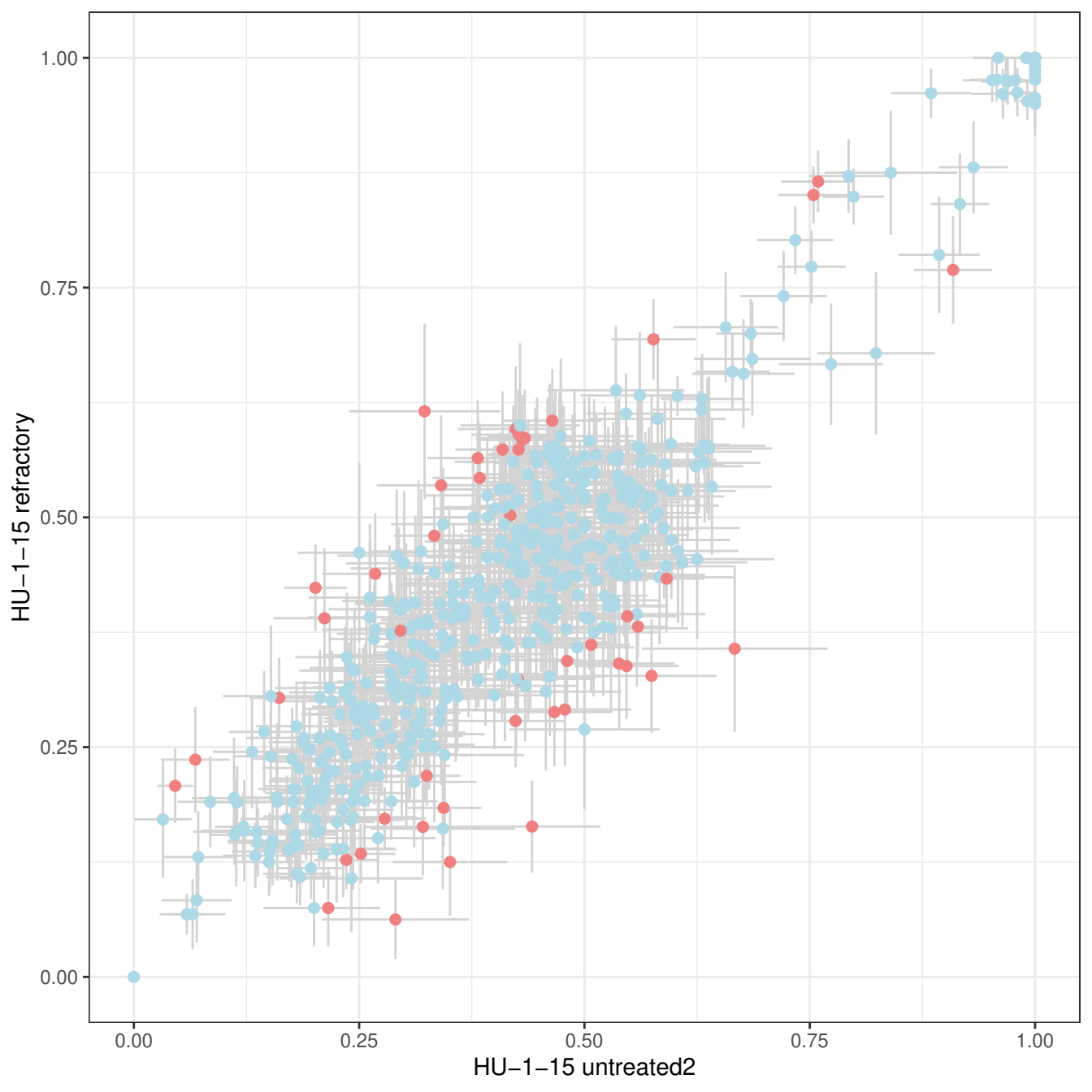

1.00
0.75
0.50
0.25
0.00
HU−1−12 untreated2
0.00
0.25
0.50
0.75
1.00
HU−1−12 untreated
ZFHX3
POT1

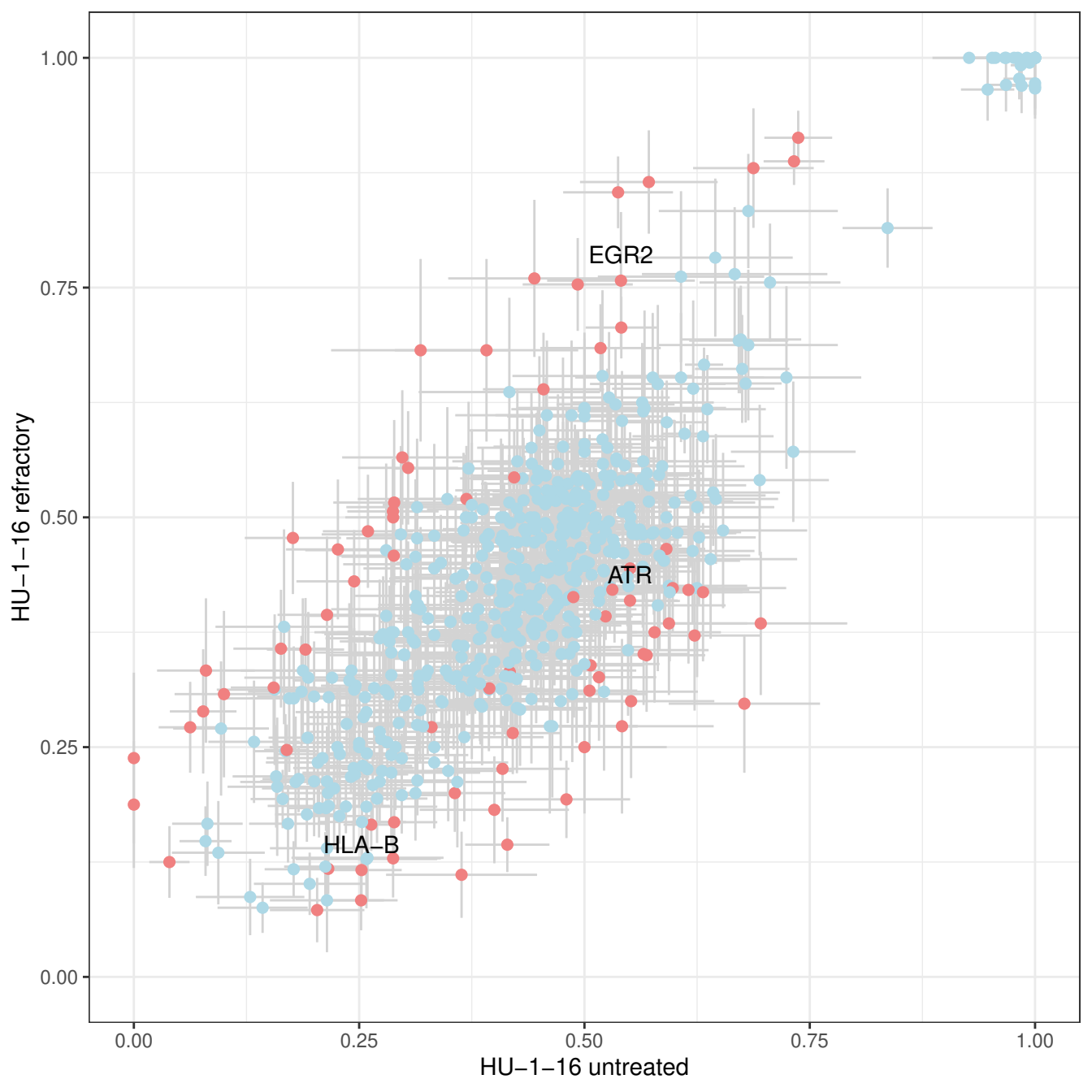

1.00
0.75
0.50
0.25
0.00
HU−1−13 relapsed
SMC3
MAP3K4
PLCB4
0.00
0.25
0.50
0.75
1.00
HU−1−13 untreated

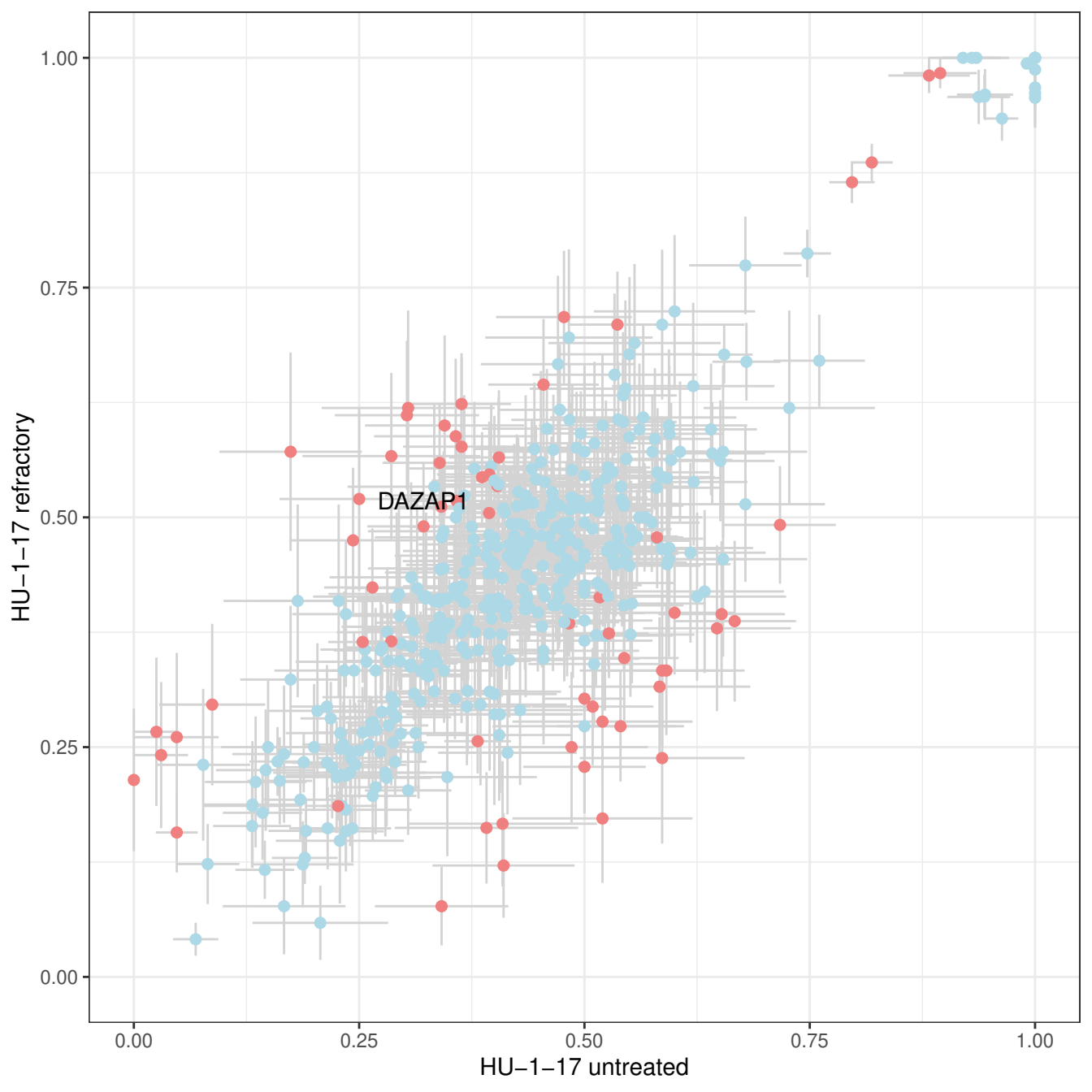

TNFAIP3
MGA
EGFR
HU−1−13 refractory
HU−1−13 relapsed
0.00
0.25
0.50
0.75
1.00

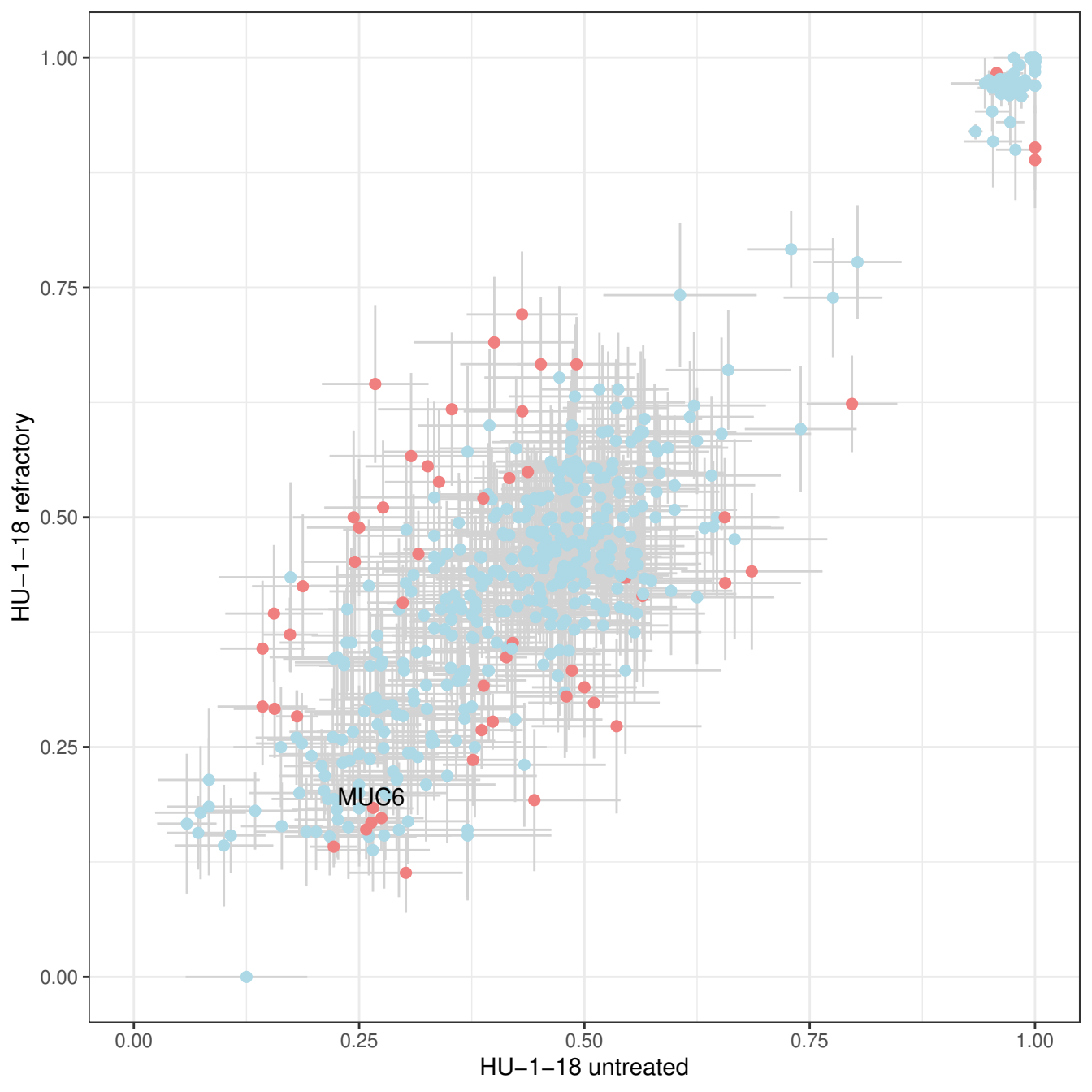

1.00
0.75
0.50
0.25
0.00
HU−1−14 relapsed
0.00
0.25
0.50
0.75
1.00
HU−1−14 untreated
TET2
snoU13
snoU13
KMT2A
GTF2I

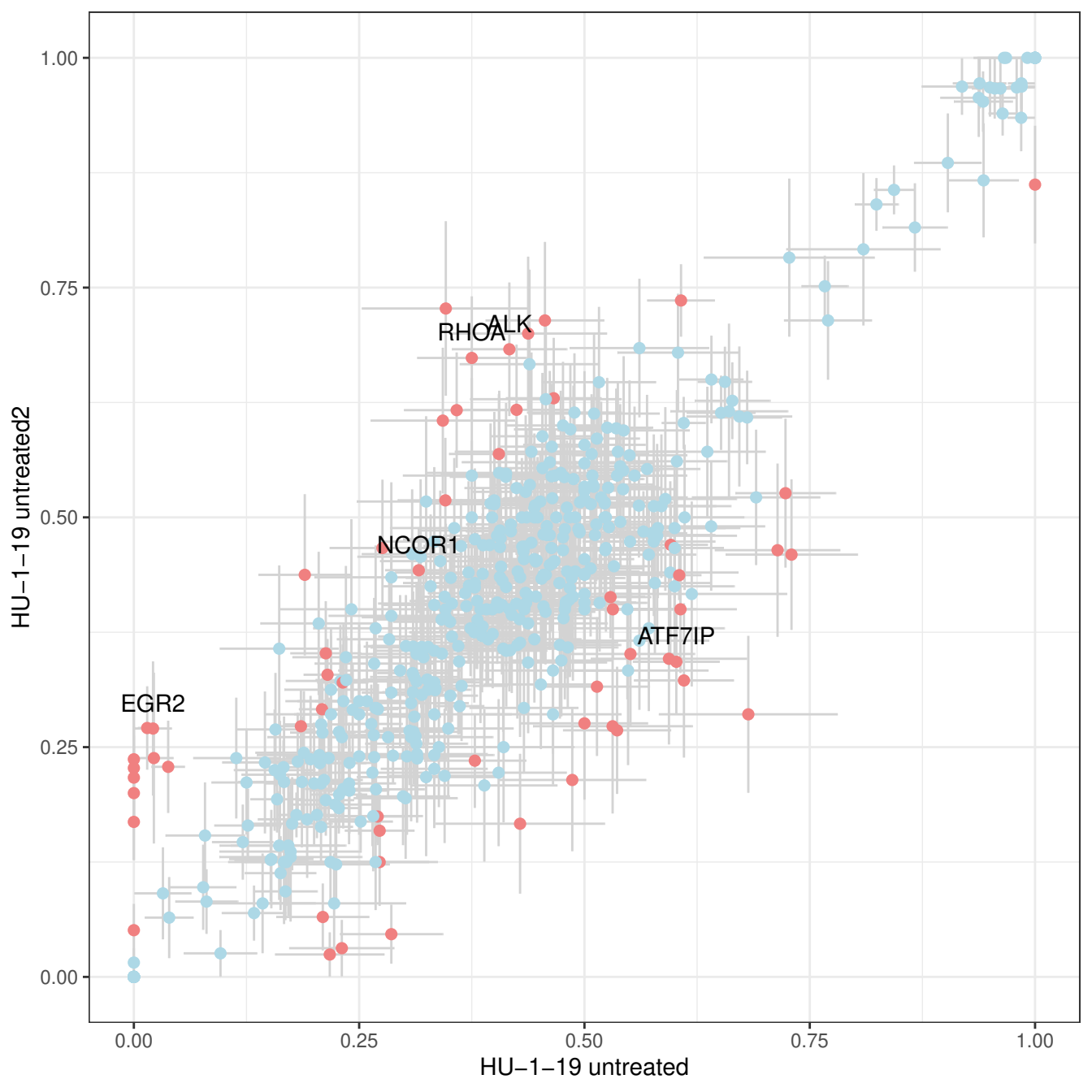

1.00
0.75
0.50
0.25
0.00
HU−1−15 untreated2
0.00
0.25
0.50
0.75
1.00
HU−1−15 untreated

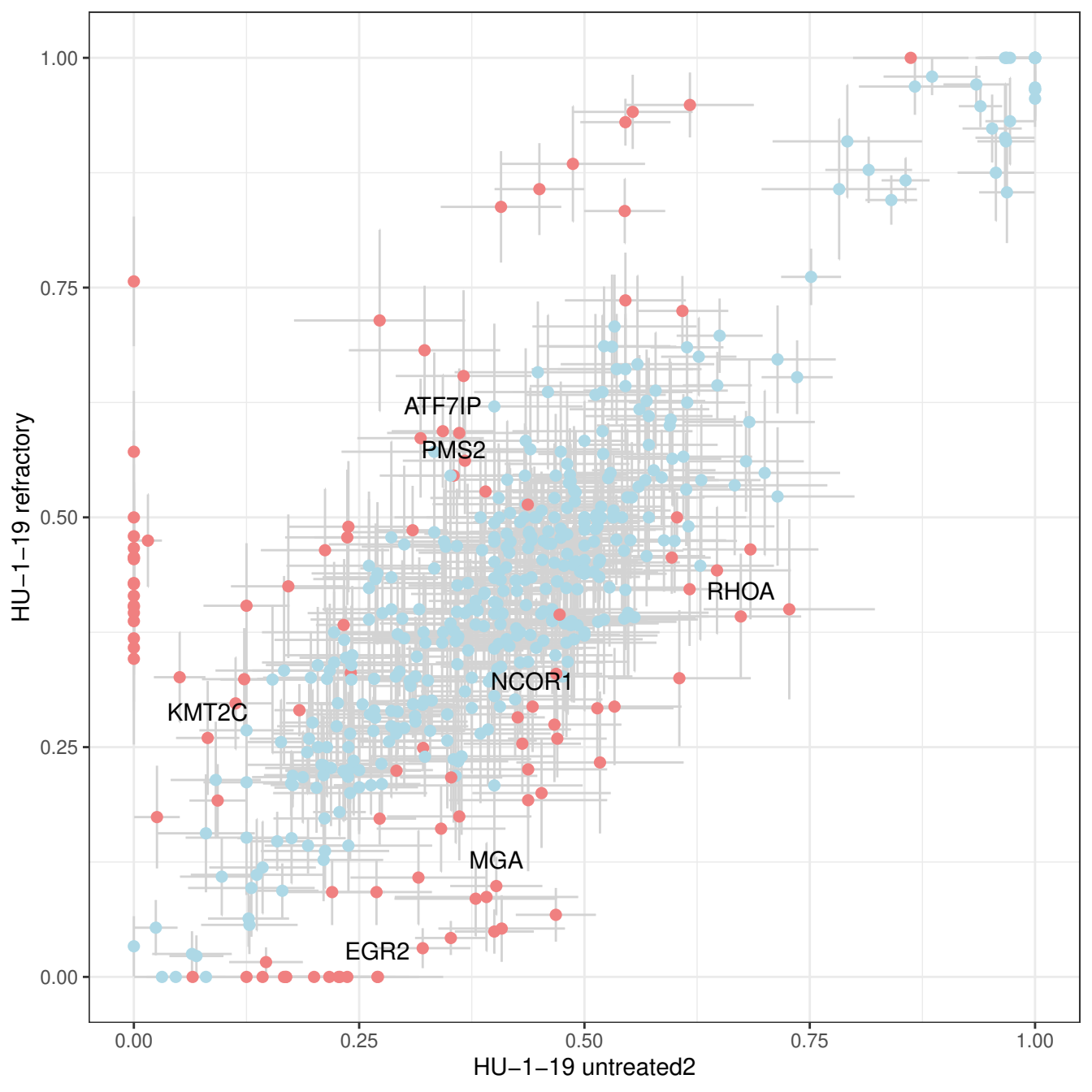

1.00
0.75
0.50
0.25
0.00
HU−1−15 refractory
0.00
0.25
0.50
0.75
1.00
HU−1−15 untreated2

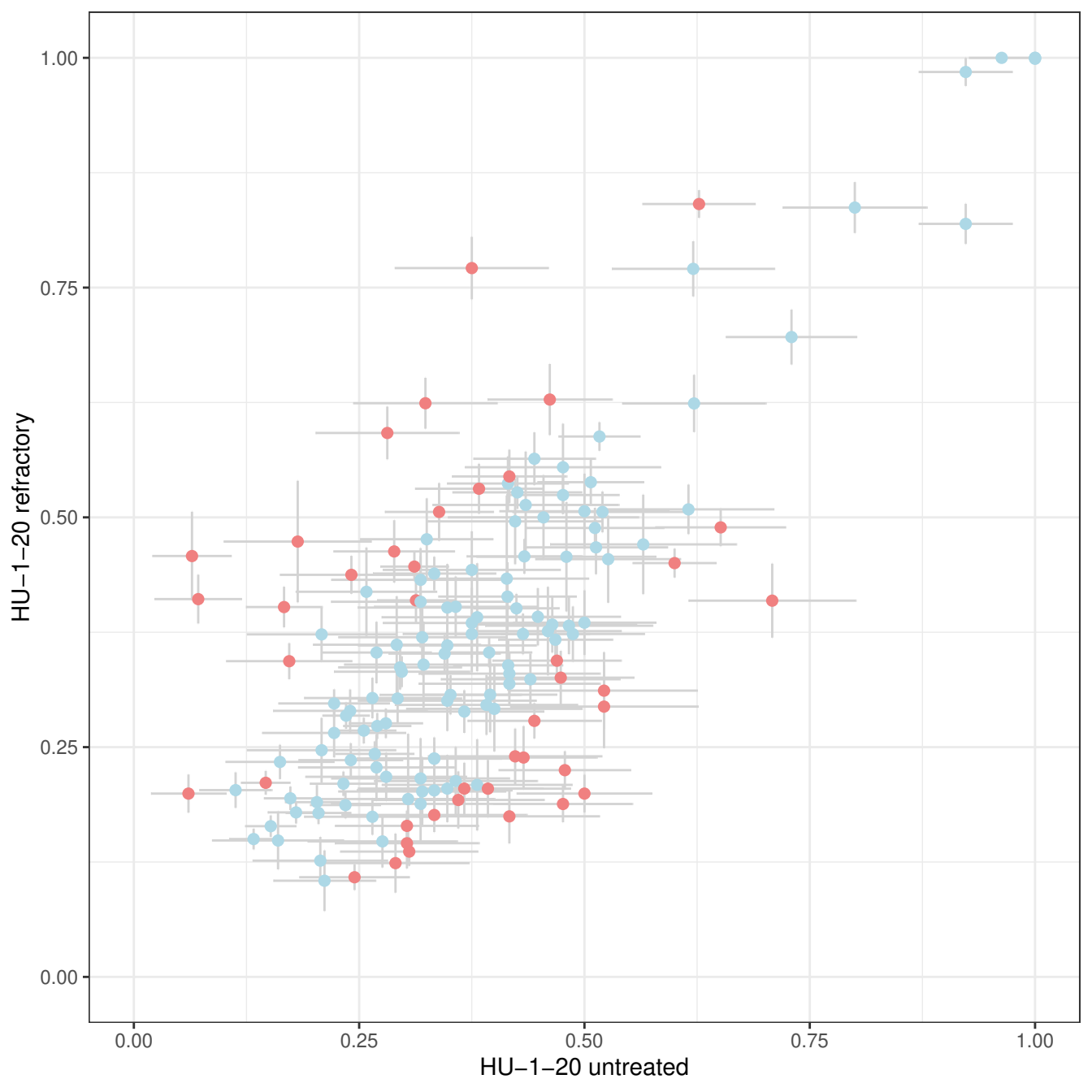

EGR2
ATR
HLA−B
HU−1−16 refractory
HU−1−16 untreated
0.00
0.25
0.50
0.75
1.00

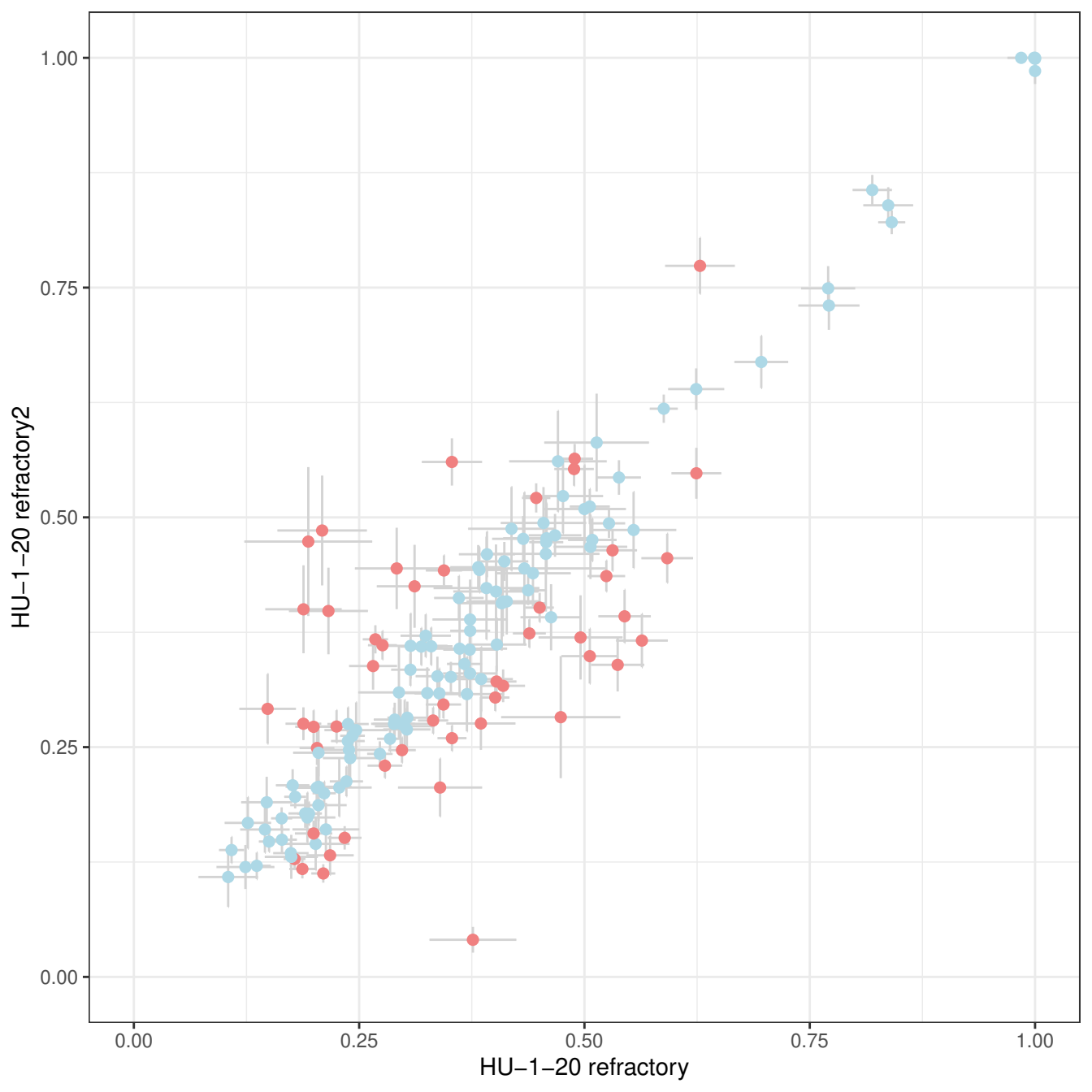

DAZAP1
HU−1−17 refractory
HU−1−17 untreated
0.00
0.25
0.50
0.75
1.00

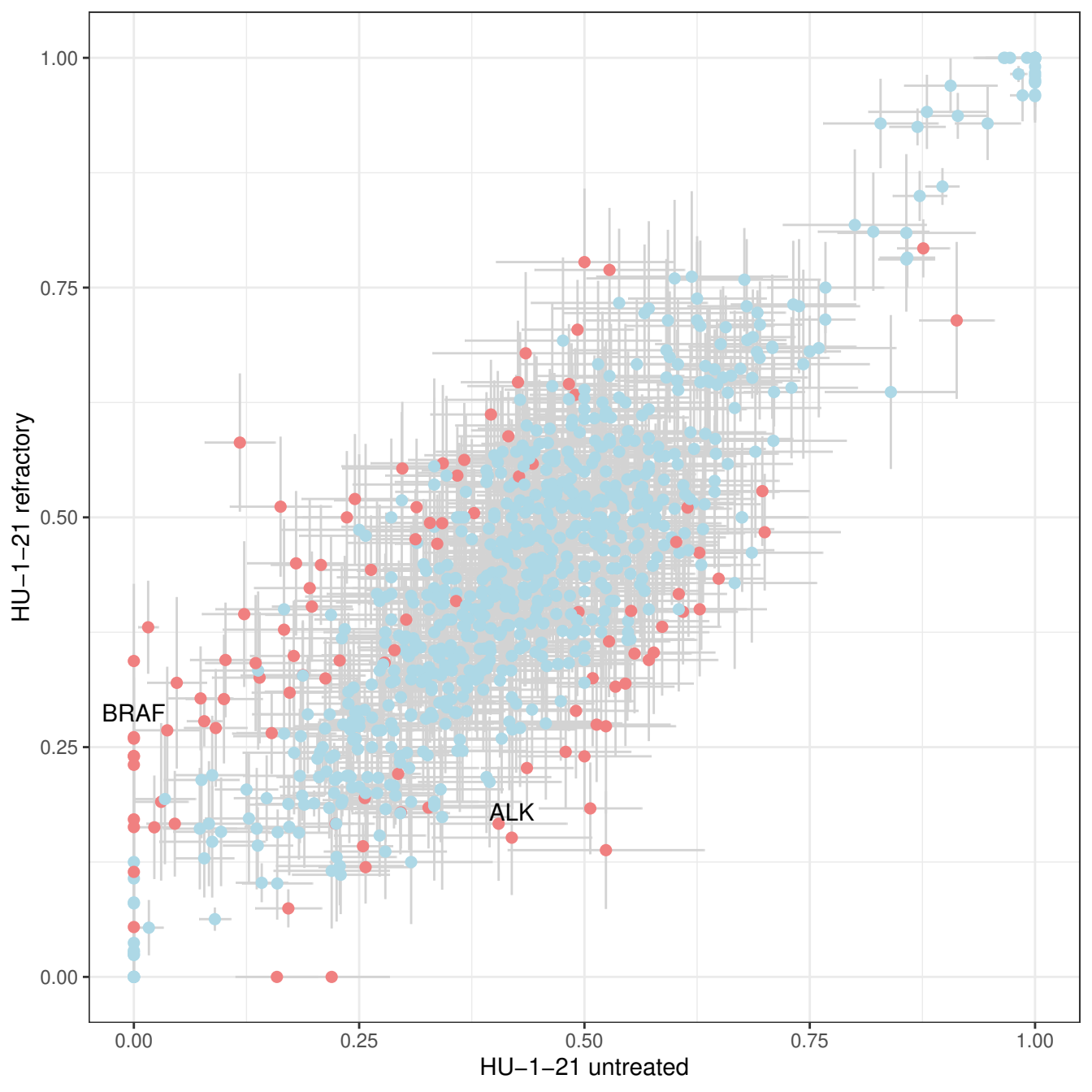

1.00
0.75
0.50
0.25
0.00
HU−1−18 refractory
0.00
0.25
0.50
0.75
1.00
HU−1−18 untreated
MUC6

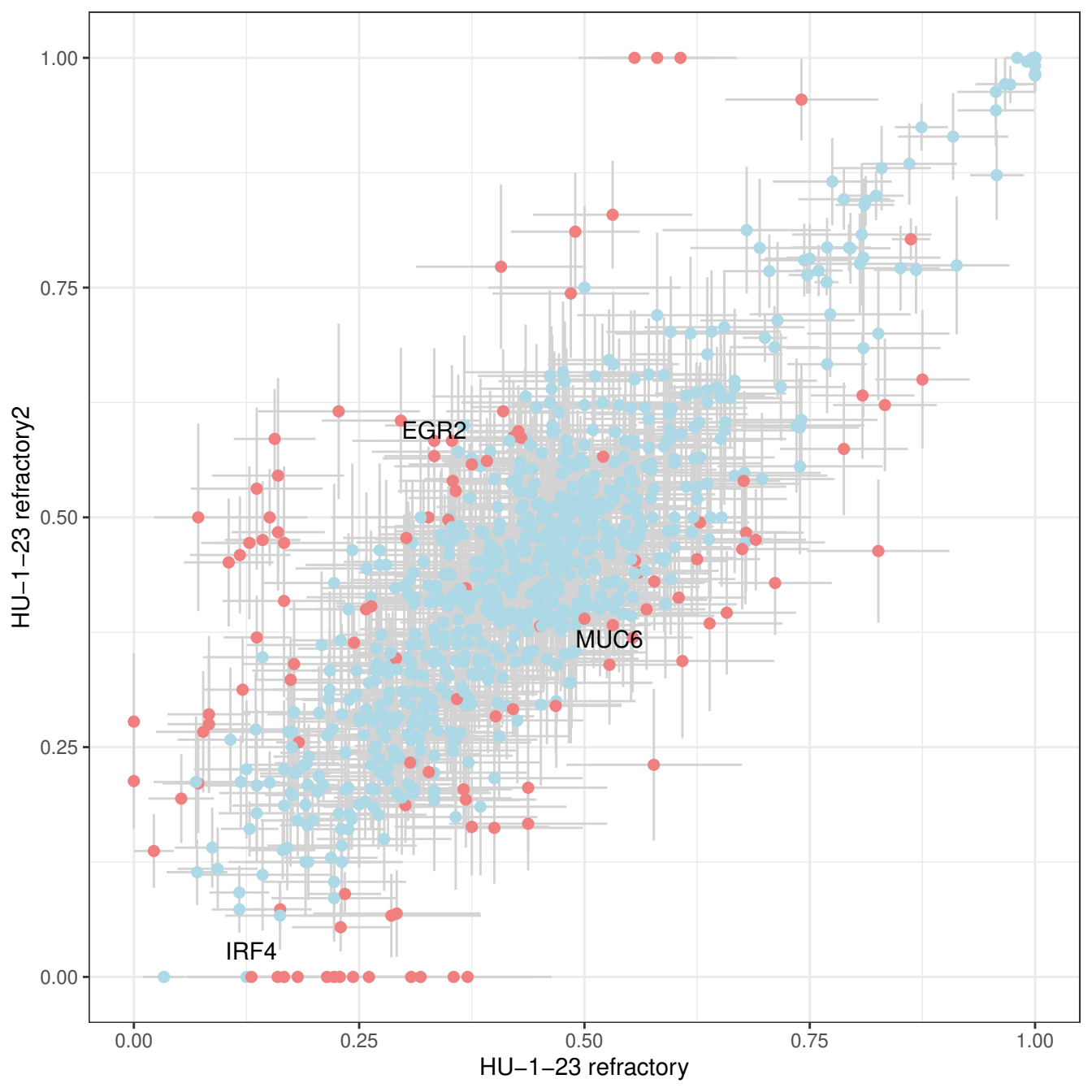

RHOA
ALK
NCOR1
ATF7IP
EGR2
1.00
0.75
0.50
0.25
0.00
HU−1−19 untreated2
0.00
0.25
0.50
0.75
1.00
HU−1−19 untreated

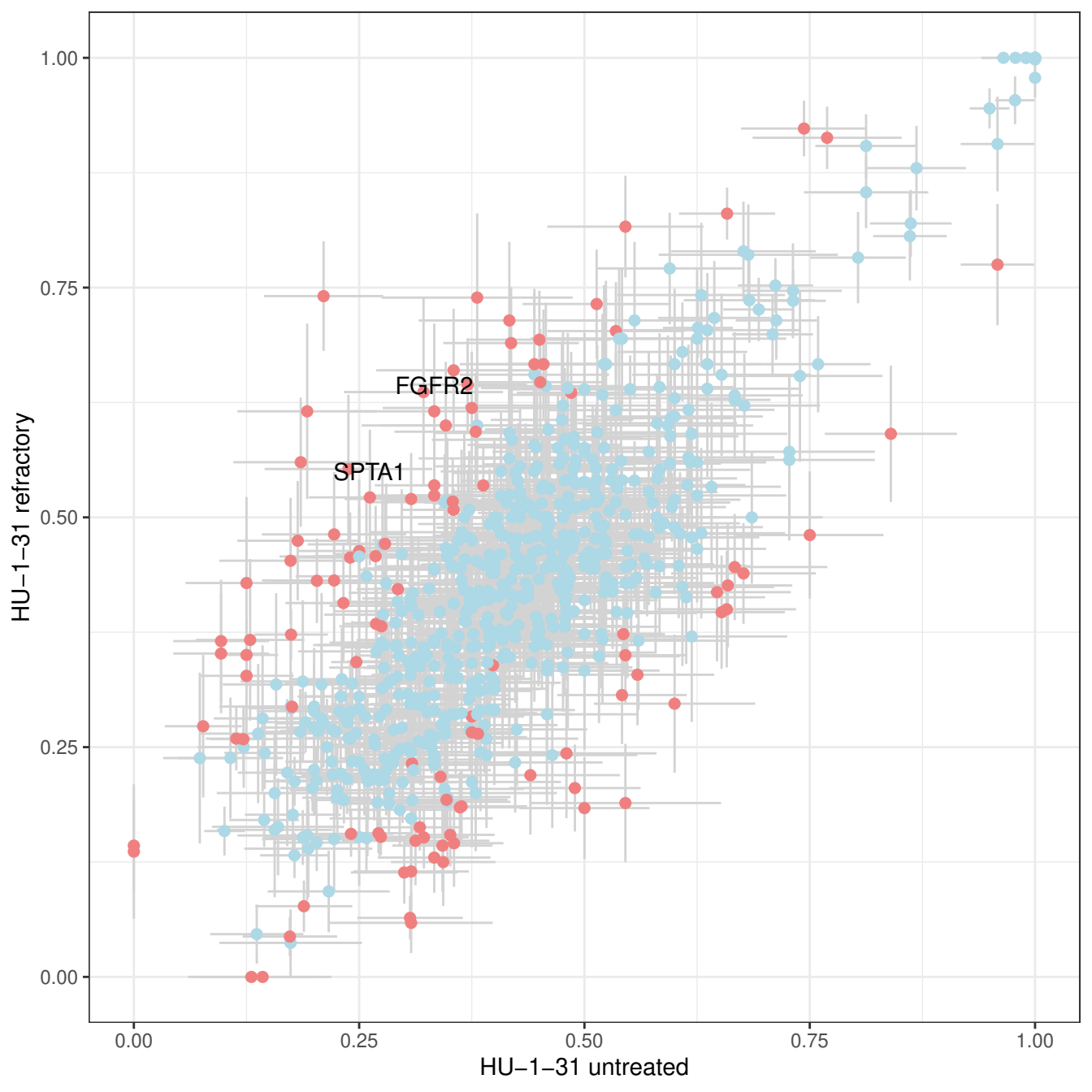

ATF7IP
PMS2
RHOA
NCOR1
KMT2C
MGA
EGR2
1.00
0.75
0.50
0.25
0.00
HU−1−19 refractory
0.00
0.25
0.50
0.75
1.00
HU−1−19 untreated2

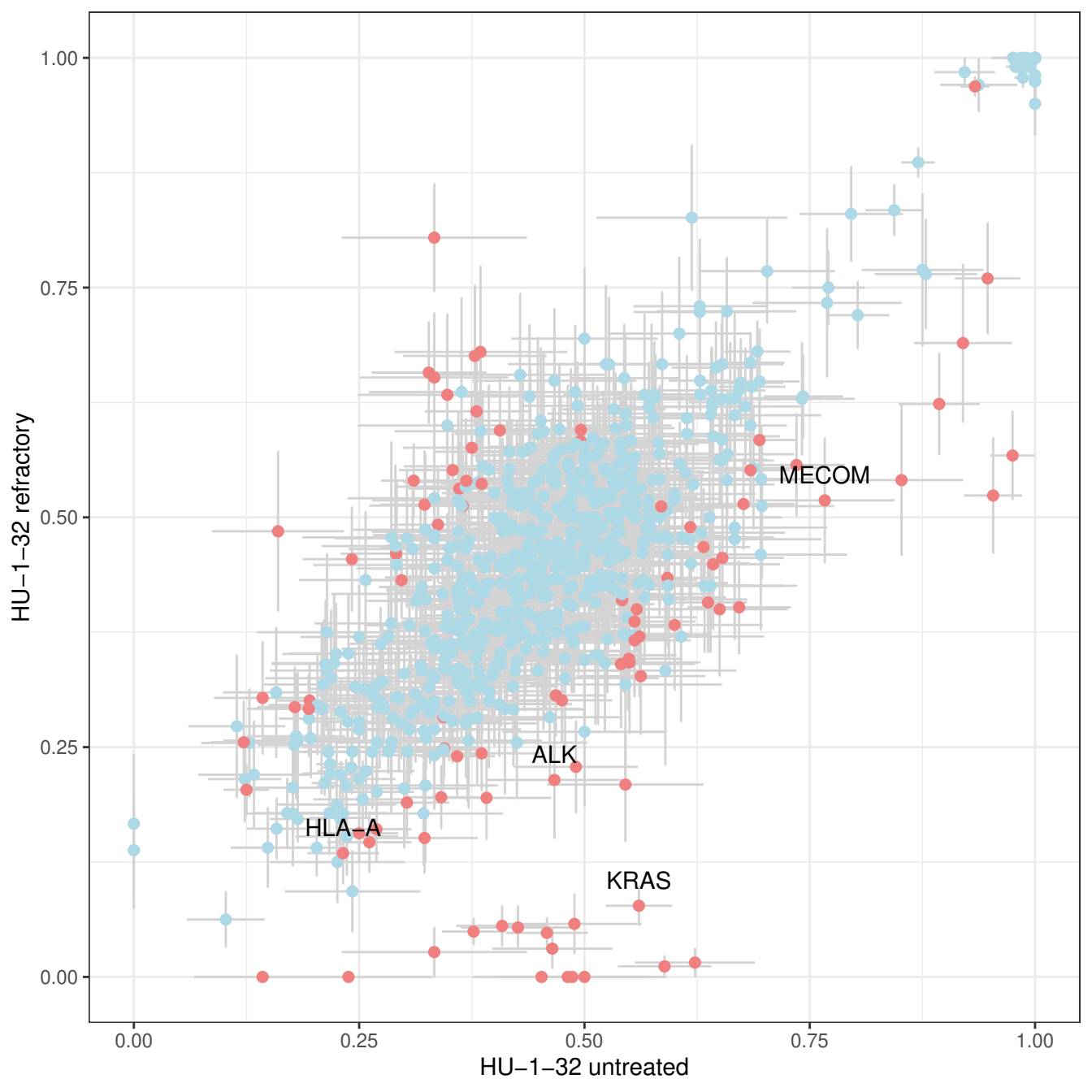

1.00
0.75
0.50
0.25
0.00
HU−1−20 refractory
0.00
0.25
0.50
0.75
1.00
HU−1−20 untreated

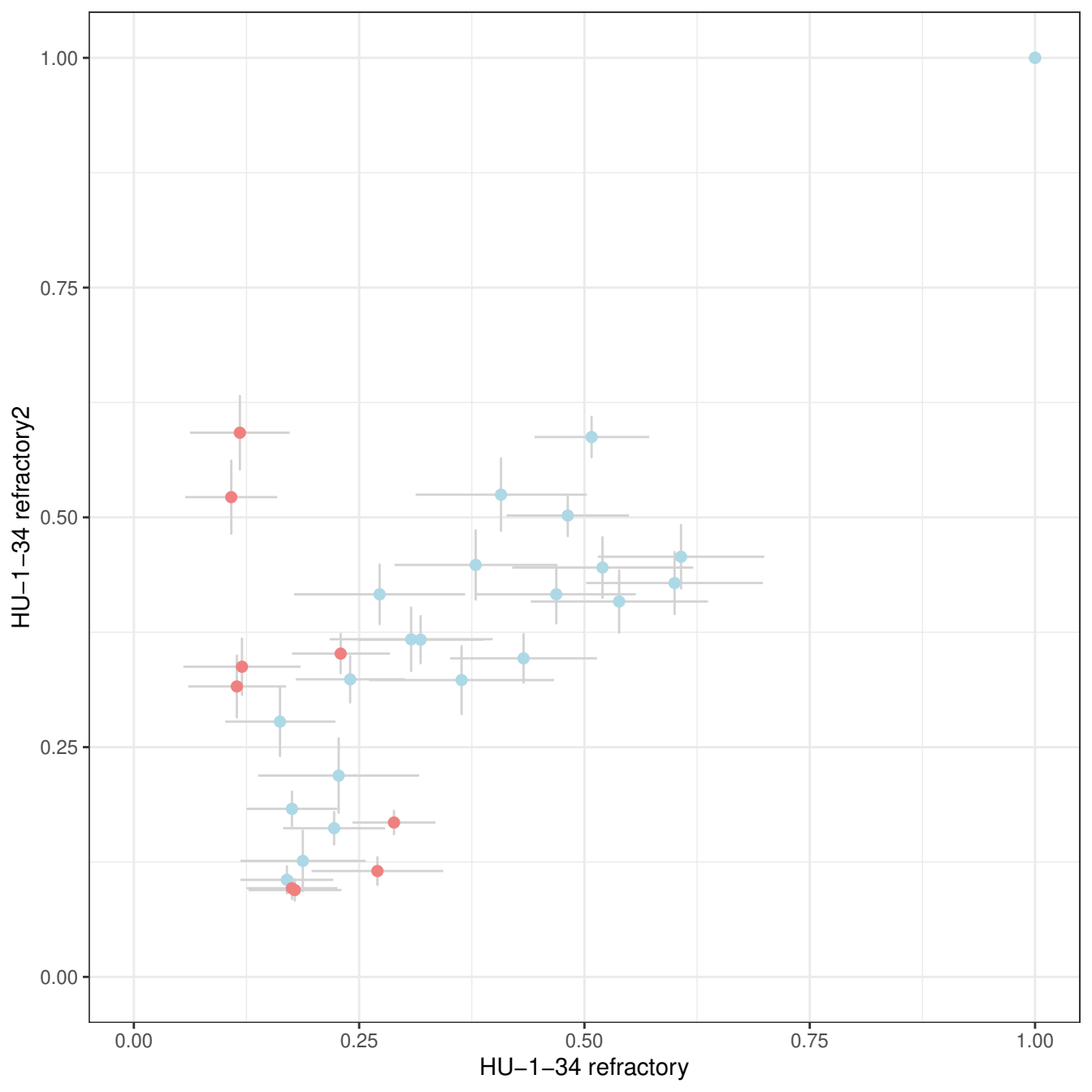

1.00
0.75
0.50
0.25
0.00
HU−1−20 refractory2
0.00
0.25
0.50
0.75
1.00
HU−1−20 refractory

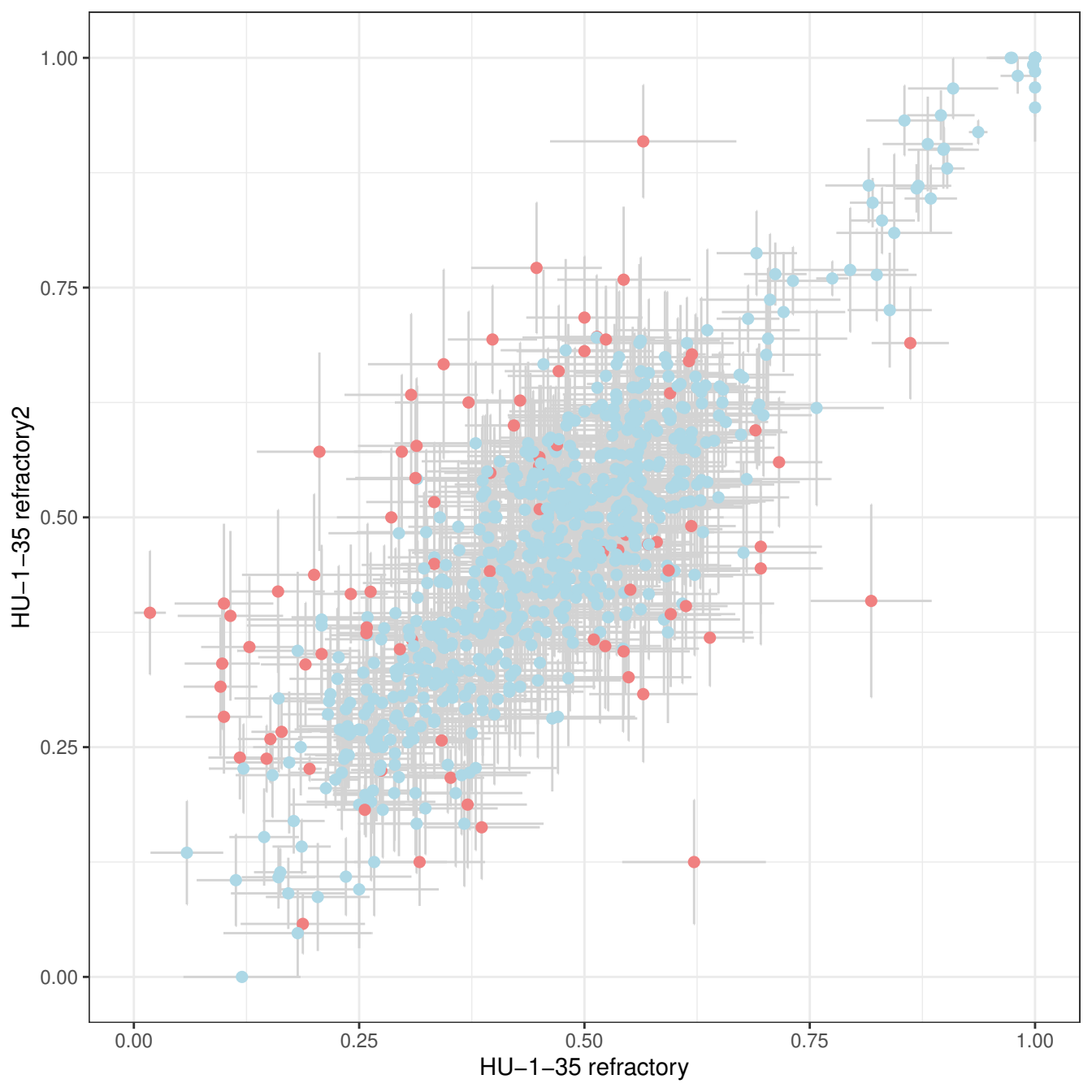

1.00
0.75
0.50
0.25
0.00
HU−1−21 refractory
BRAF
ALK
0.00
0.25
0.50
0.75
1.00
HU−1−21 untreated

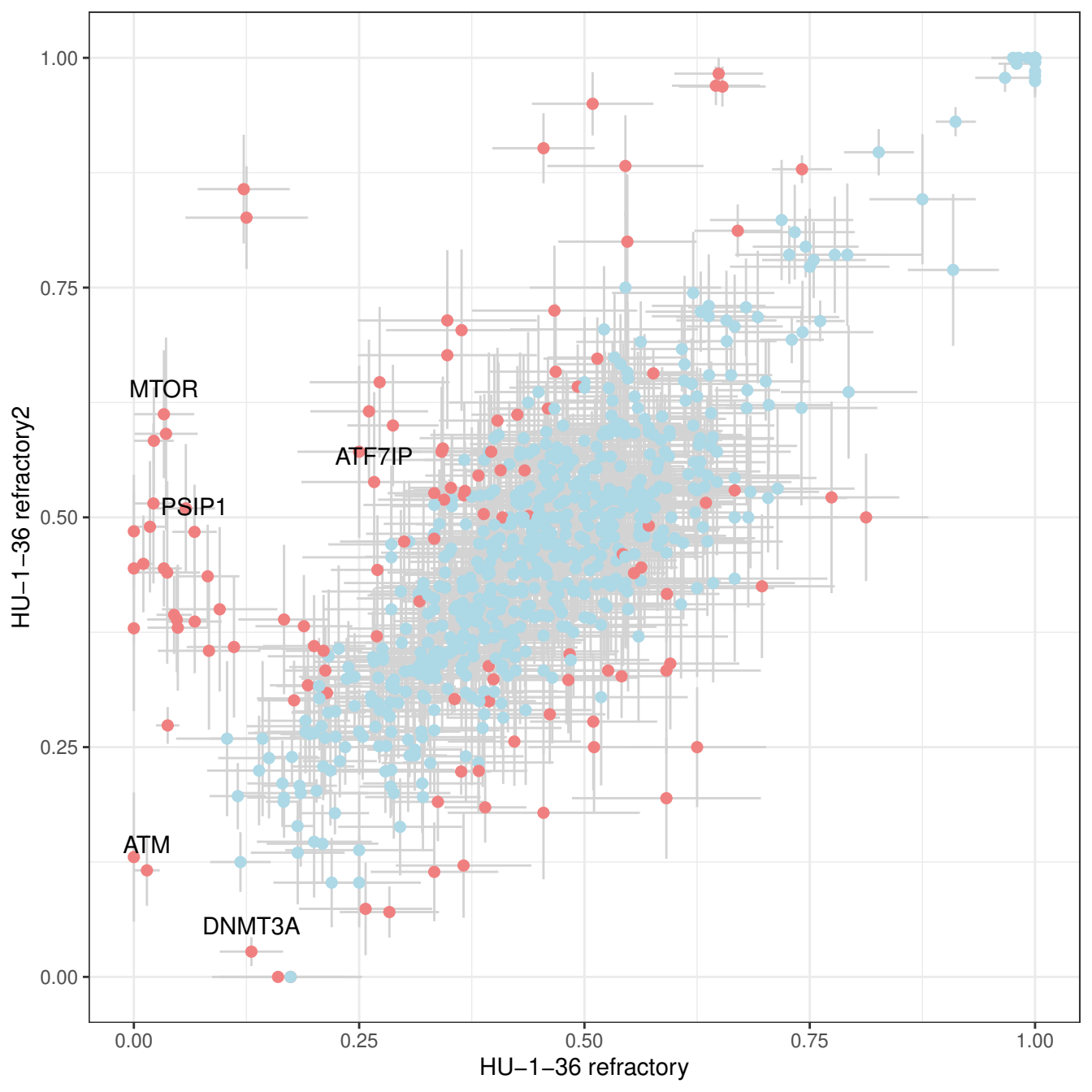

EGR2
MUC6
IRF4
1.00
0.75
0.50
0.25
0.00
HU−1−23 refractory2
0.00
0.25
0.50
0.75
1.00
HU−1−23 refractory

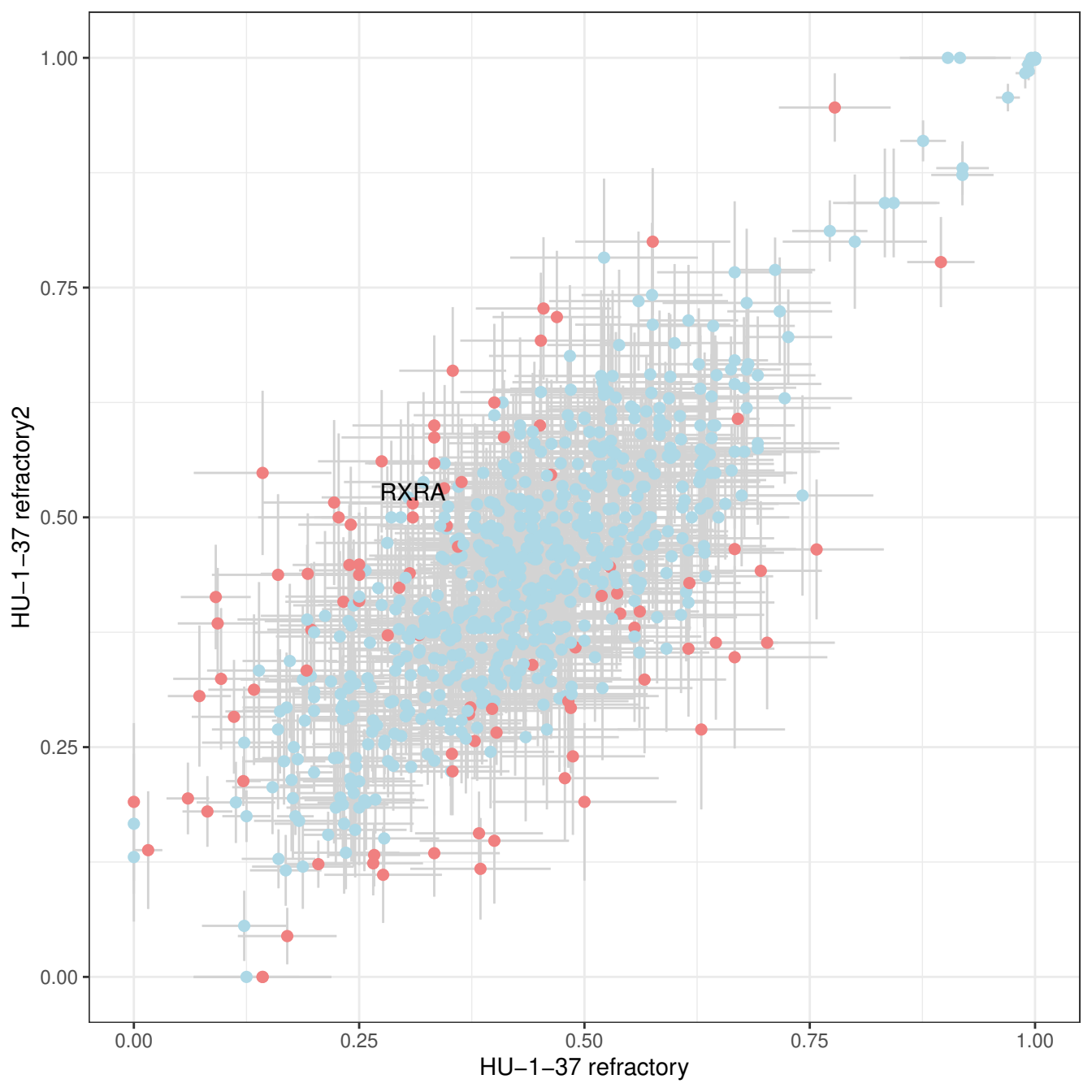

FGFR2
SPTA1
HU−1−31 refractory
HU−1−31 untreated
0.00
0.25
0.50
0.75
1.00


1.00
0.75
0.50
0.25
0.00
HU−1−32 refractory
0.00
0.25
0.50
0.75
1.00
HU−1−32 untreated
MECOM
ALK
HLA−A
KRAS

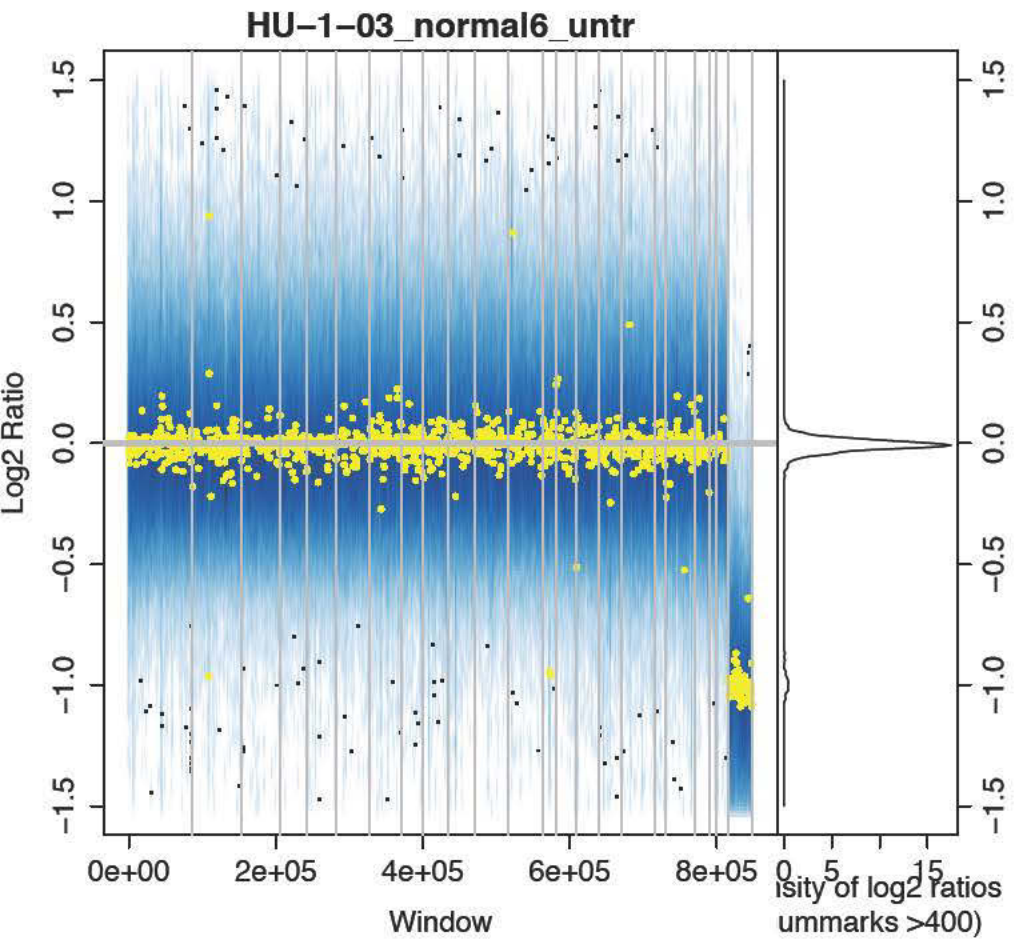

1.00
0.75
0.50
0.25
0.00
HU−1−34 refractory2
0.00
0.25
0.50
0.75
1.00
HU−1−34 refractory

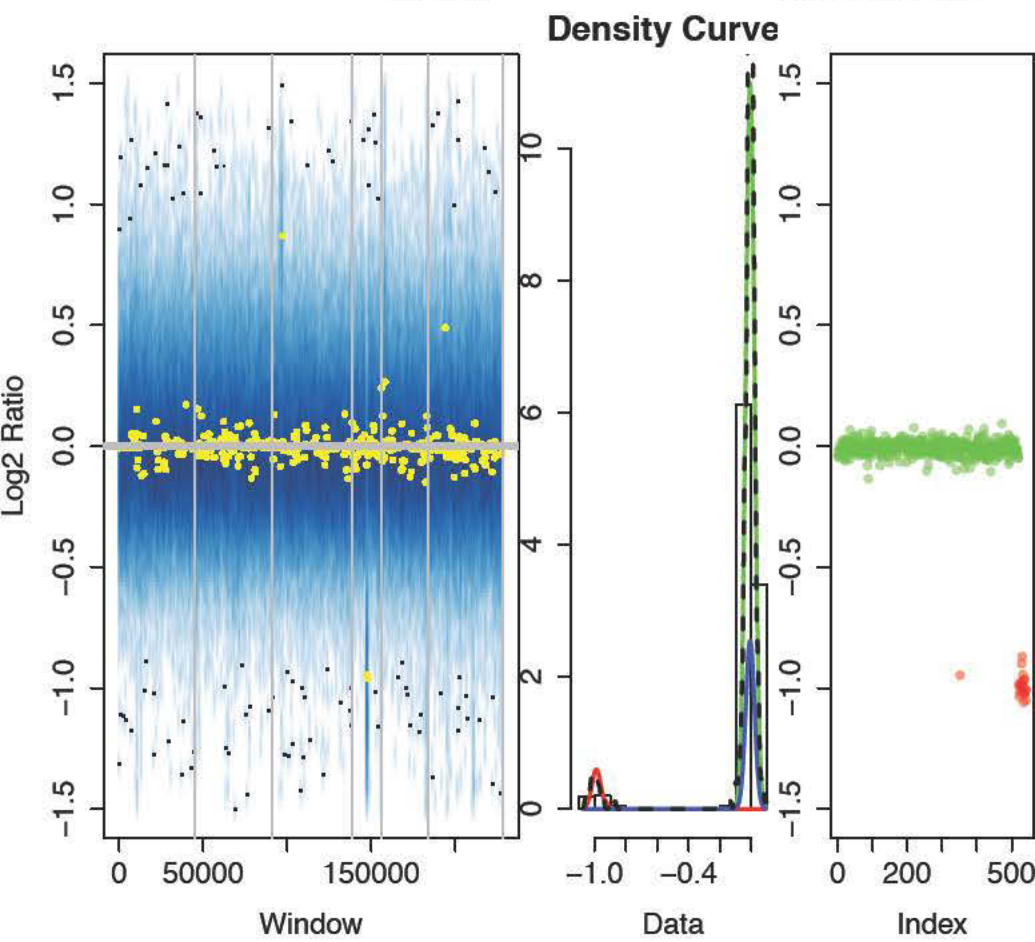

1.00
0.75
0.50
0.25
0.00
HU−1−35 refractory2
0.00
0.25
0.50
0.75
1.00
HU−1−35 refractory

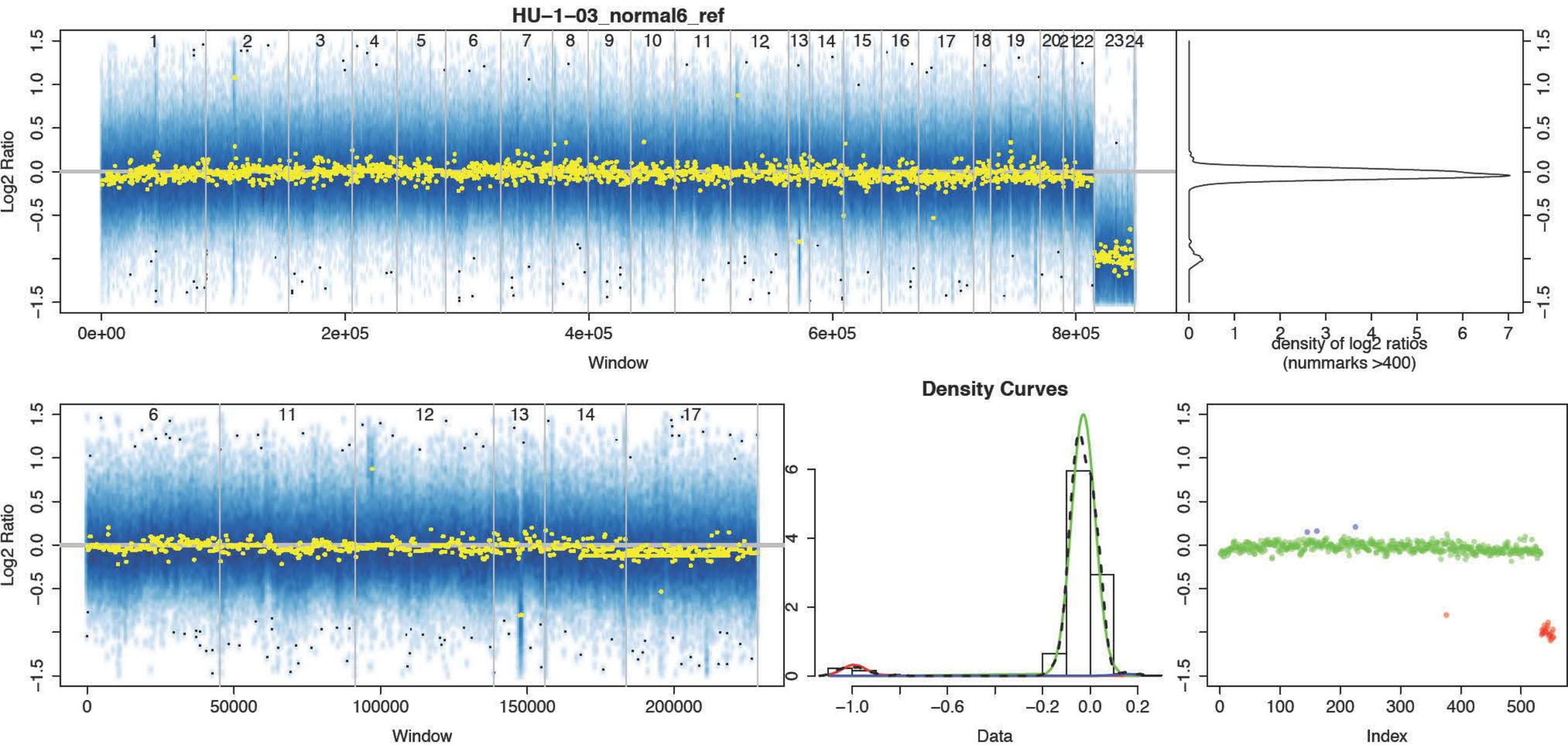

MTOR
ATF7IP
PSIP1
ATM
DNMT3A
1.00
0.75
0.50
0.25
0.00
0.00
0.25
0.50
0.75
1.00
HU−1−36 refractory2
HU−1−36 refractory

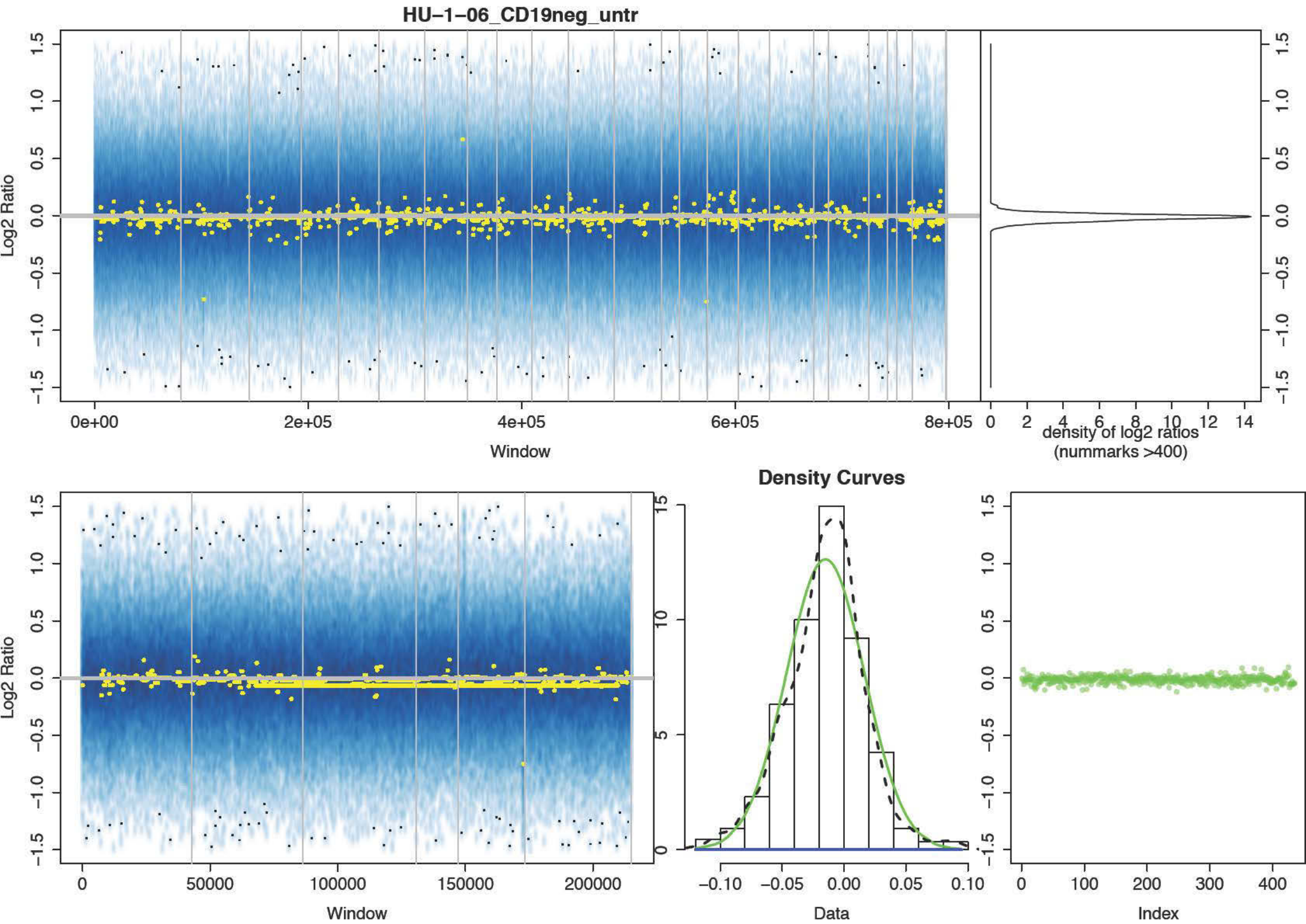

RXRA
HU−1−37 refractory
HU−1−37 refractory2
0.00
0.25
0.50
0.75
1.00

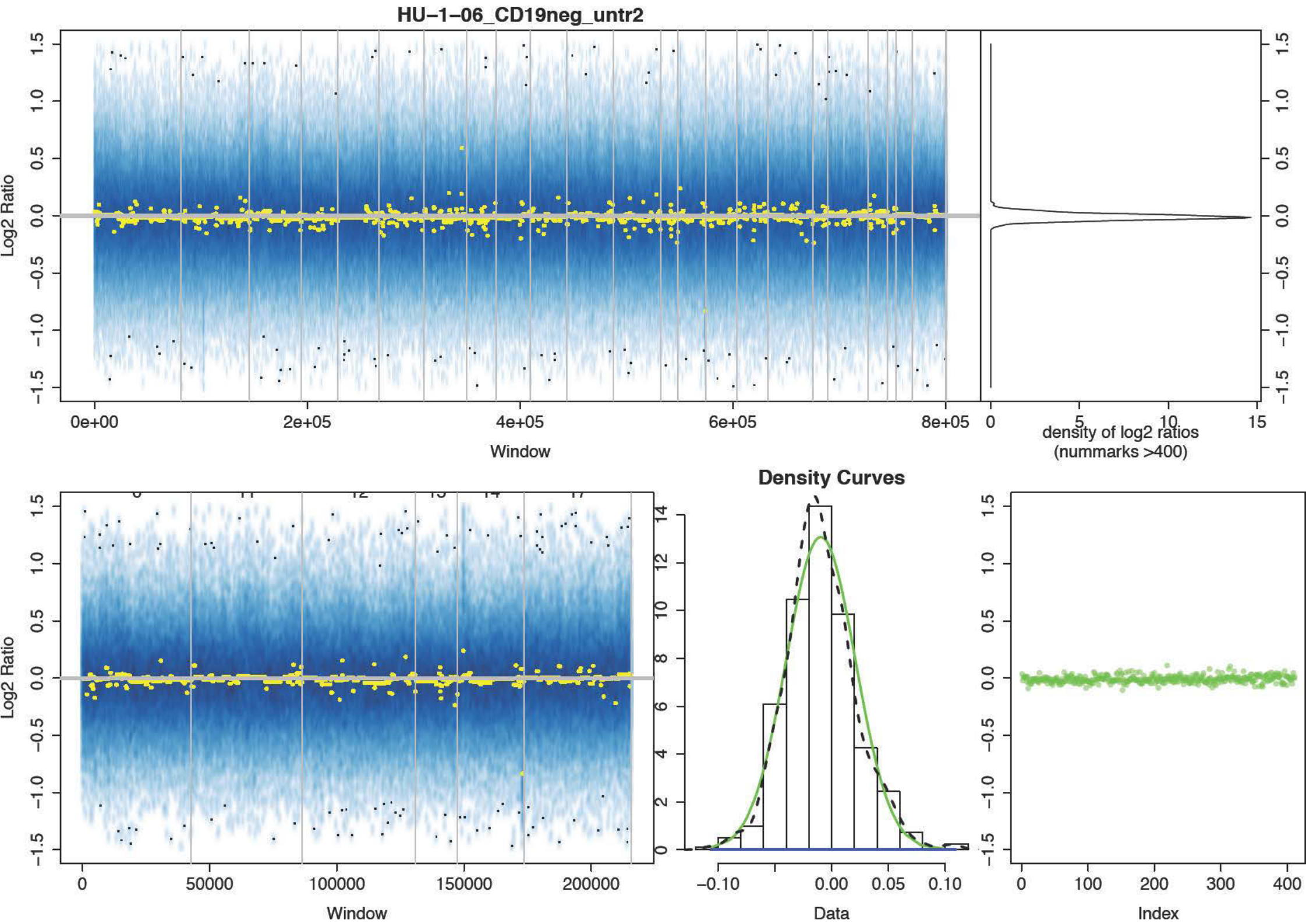
Suppl. figure 3: Copy number estimation based on read counts from exome sequencing. First panel genome wide log2 ratio of tumor vs. control counts. Second panel copy number log2 ratio plot for regions analyzed by FISH. Third panel presenting density of log2 ratio regions. Fourth panel presenting genome wide log2 copy number.

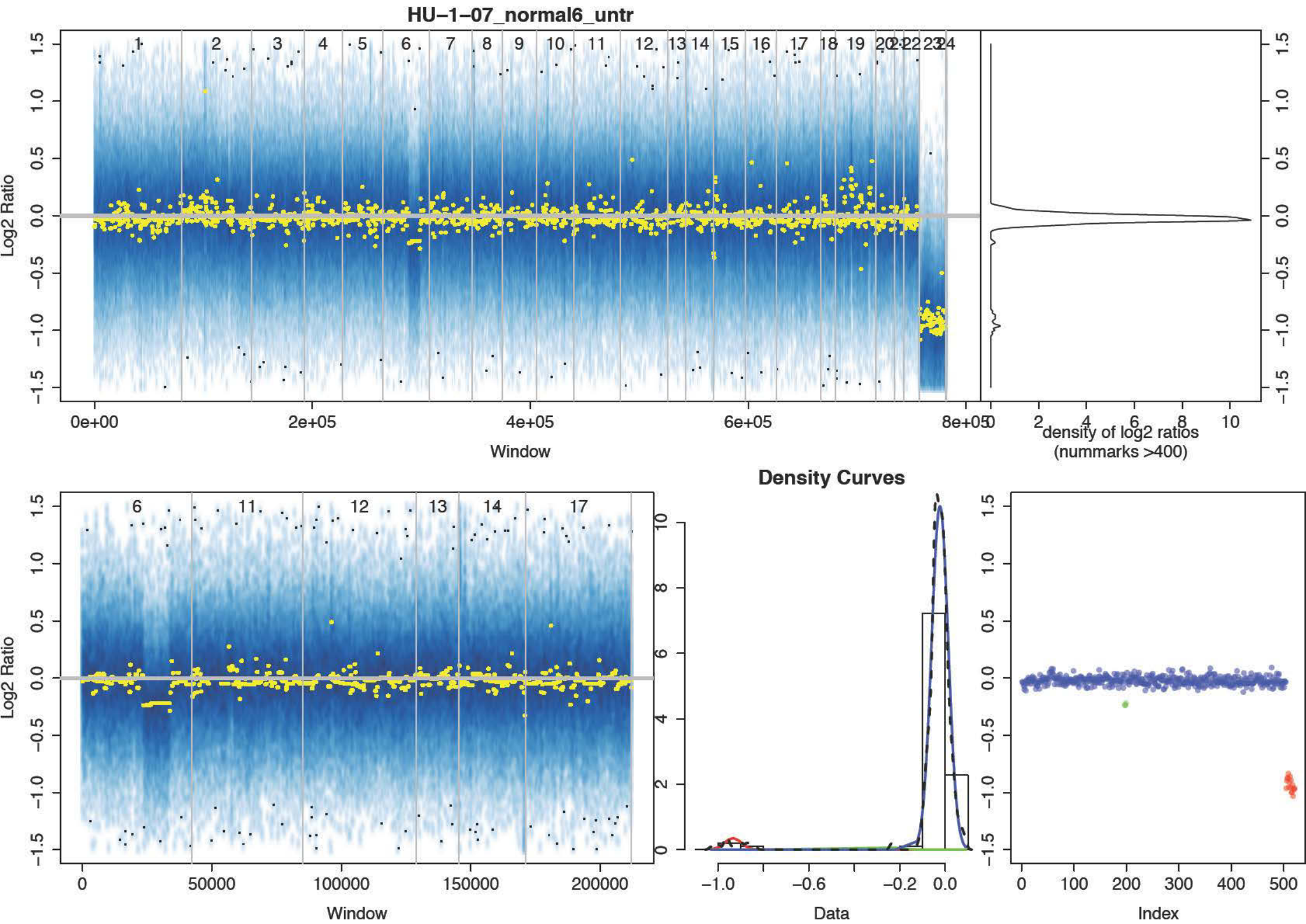

HU−1−03_normal6_untr
Log2 Ratio
1.5
1.0
0.5
0.0
−0.5
−1.0
−1.5
0e+00
2e+05
4e+05
6e+05
8e+05
Window
0
5
15
sity of log2 ratios
ummarks >400)

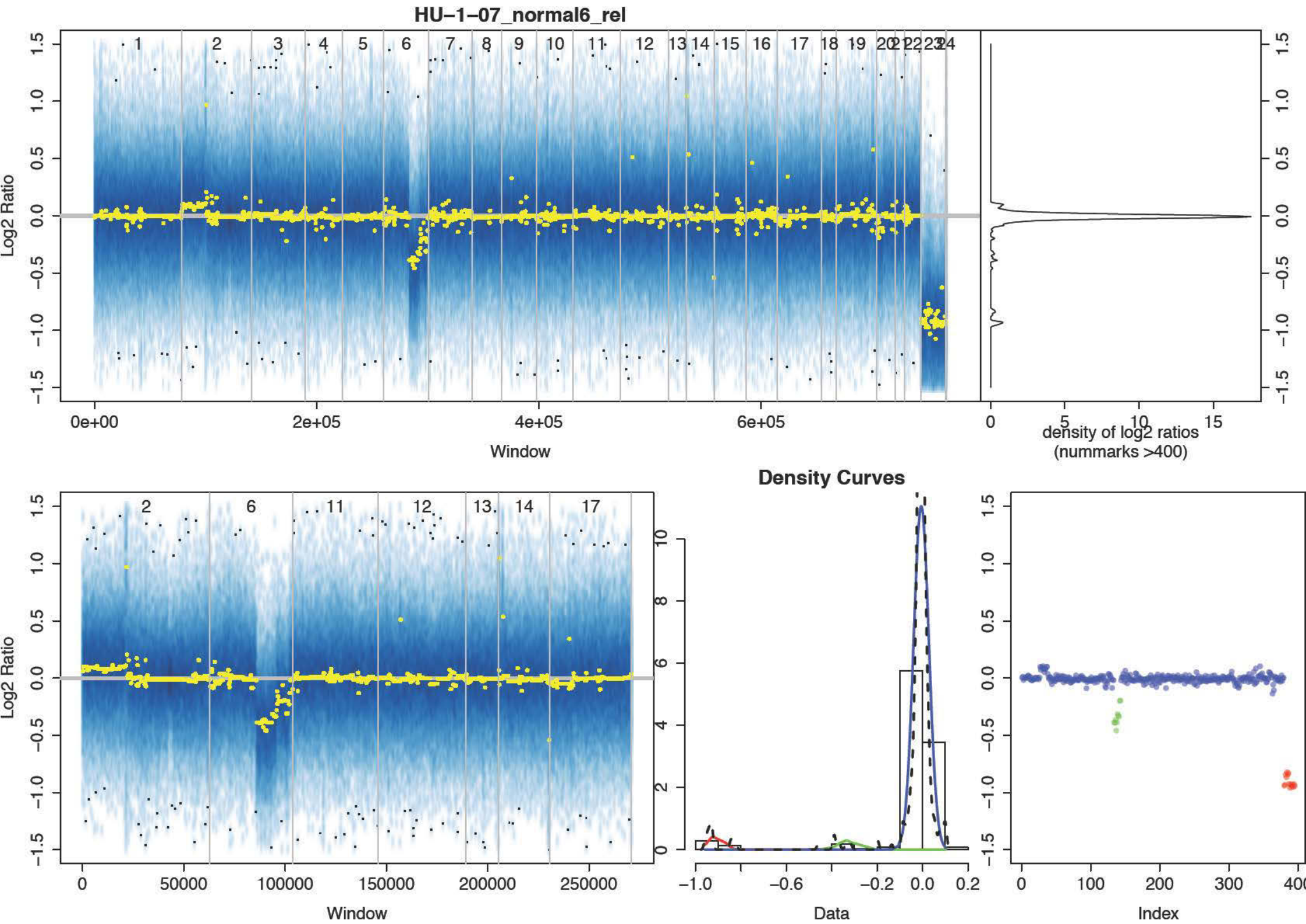

Density Curve
Log2 Ratio
1.5
1.0
0.5
0.0
−0.5
−1.0
−1.5
0
50000
150000
Window
0
2
4
6
8
10
−1.0
−0.4
Data
0
200
500
Index

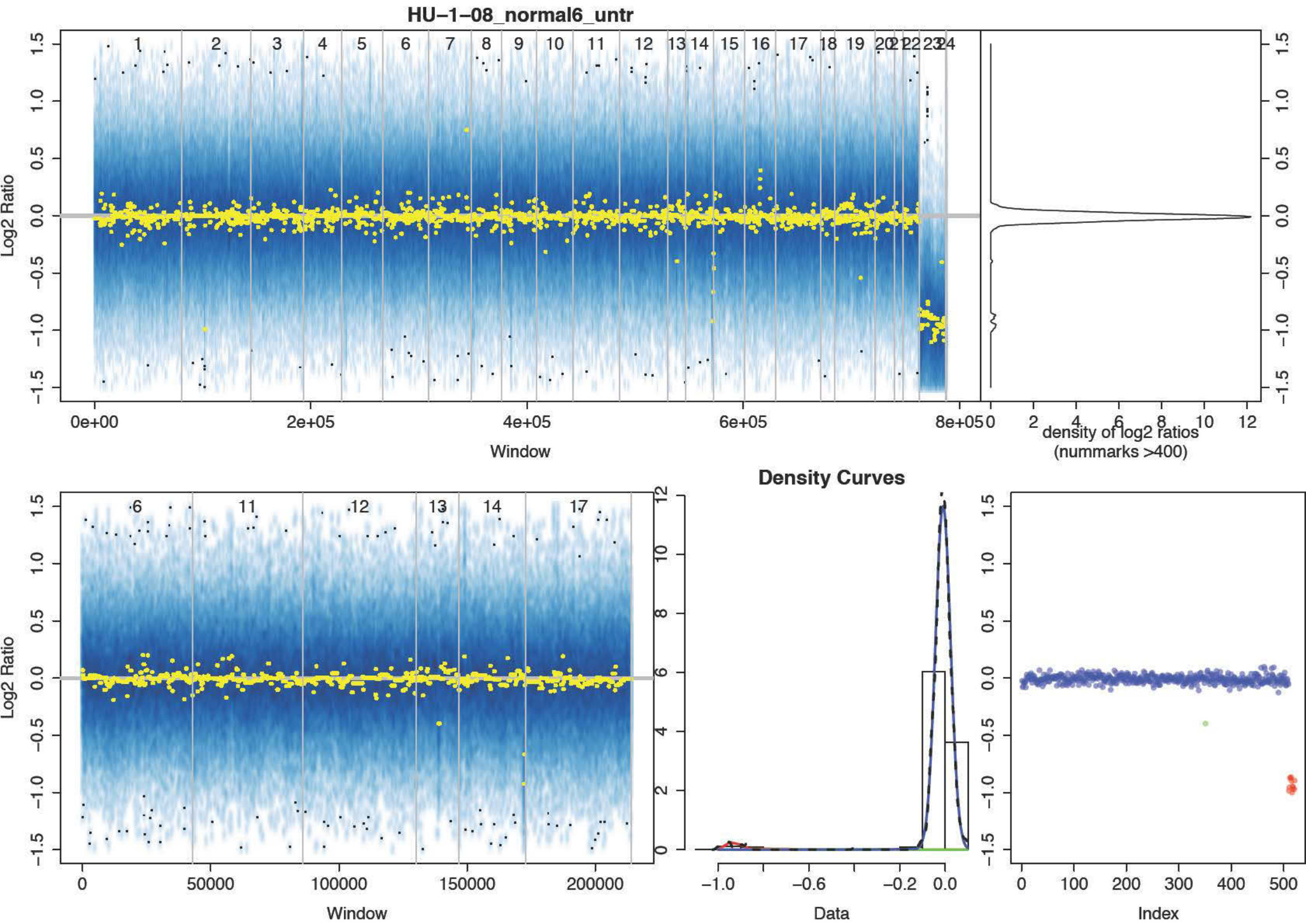

HU−1−03_normal6_ref
Log2 Ratio
Window
density of log2 ratios
(nummarks >400)
Density Curves
Data
Index

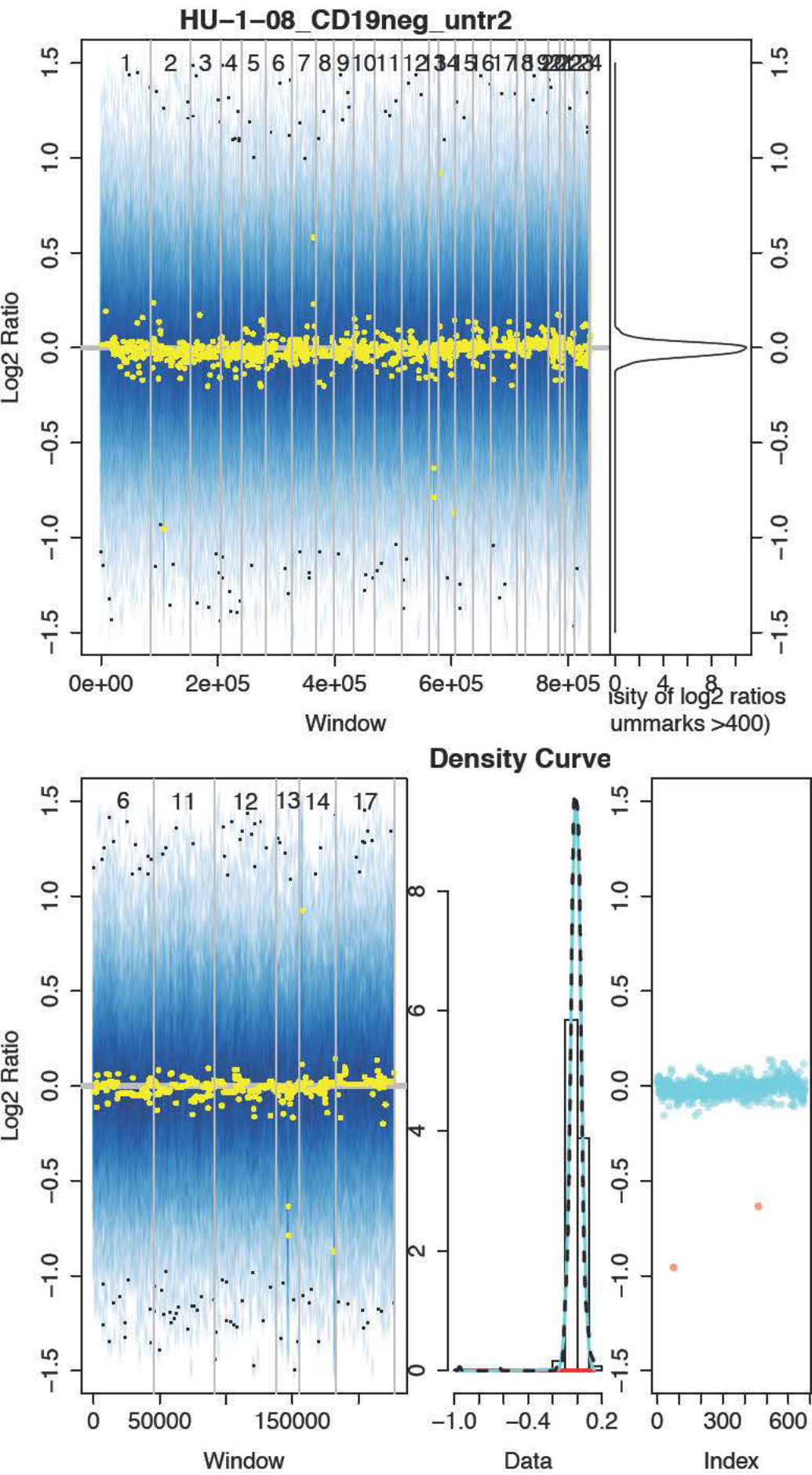

HU−1−06_CD19neg_untr
Log2 Ratio
Window
density of log2 ratios
(nummarks >400)
Density Curves
Log2 Ratio
Window
Data
Index

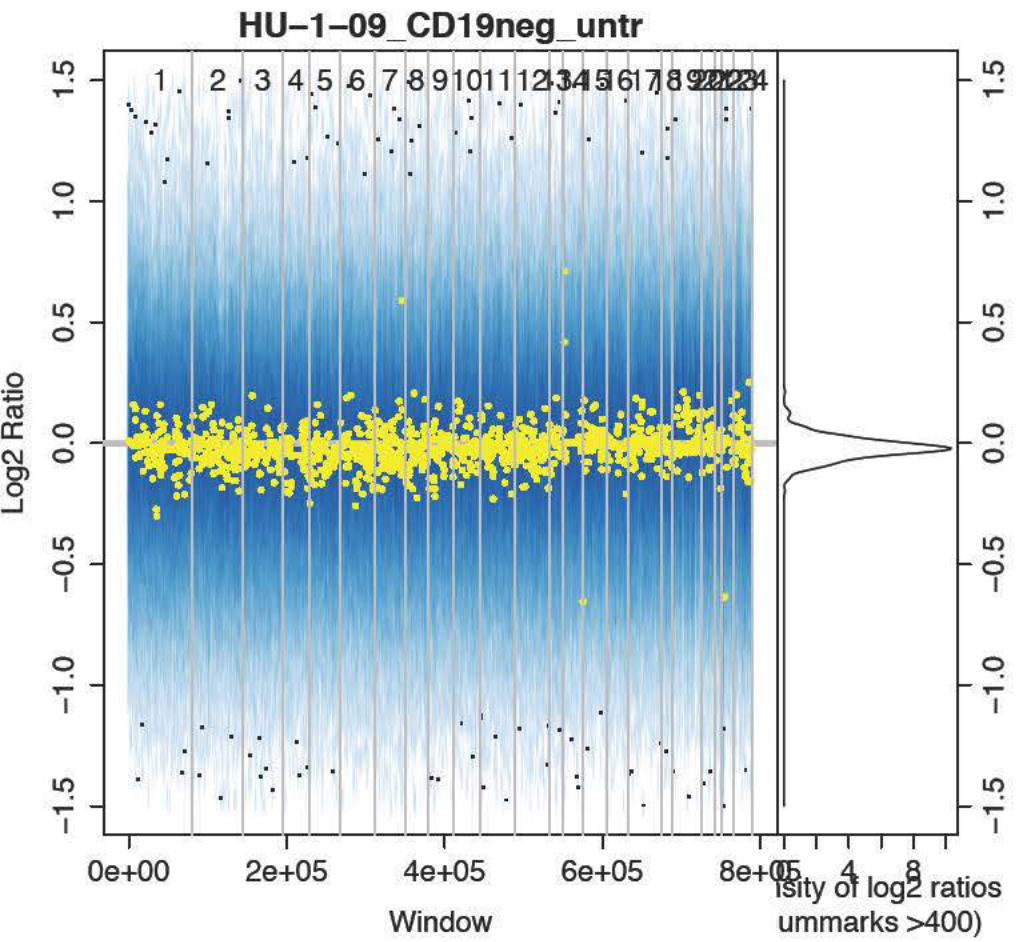

HU−1−06_CD19neg_untr2
Log2 Ratio
Window
density of log2 ratios
(nummarks >400)
Density Curves
Log2 Ratio
Window
Data
Index

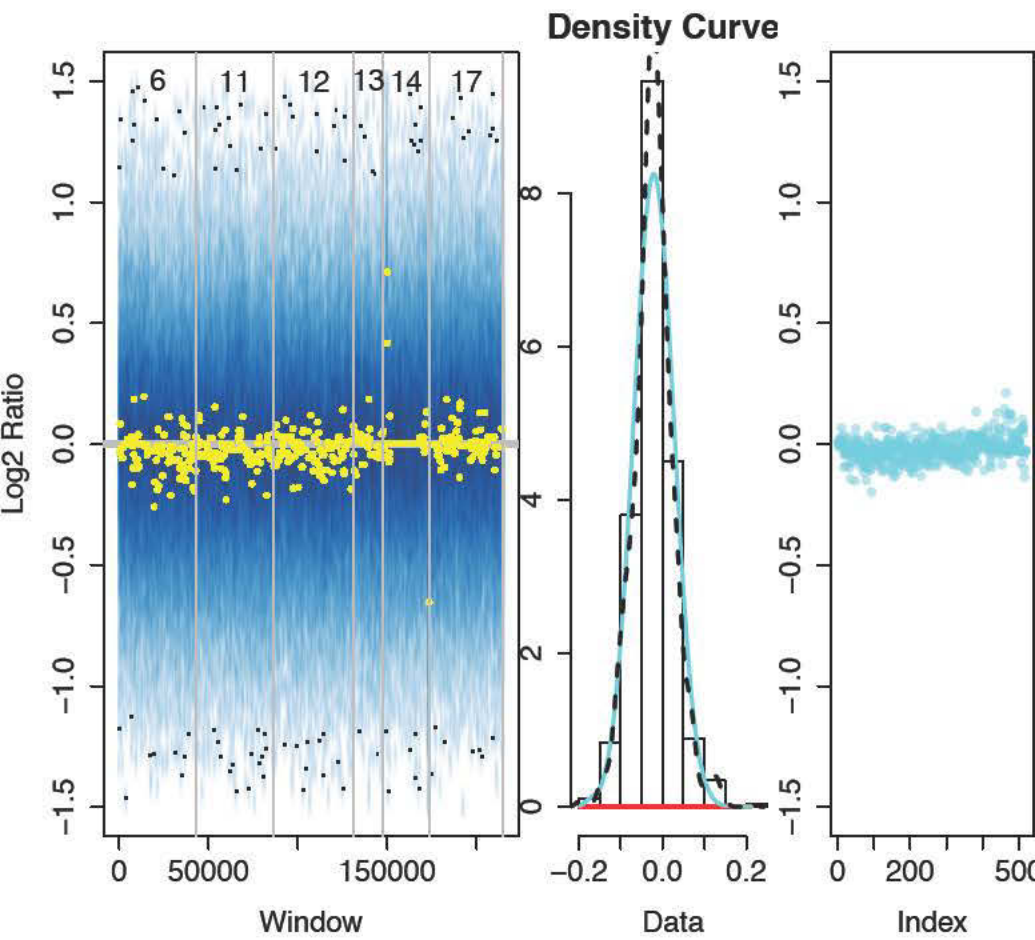

HU−1−07_normal6_untr
Log2 Ratio
Window
density of log2 ratios
(nummarks >400)
Density Curves
Data
Index

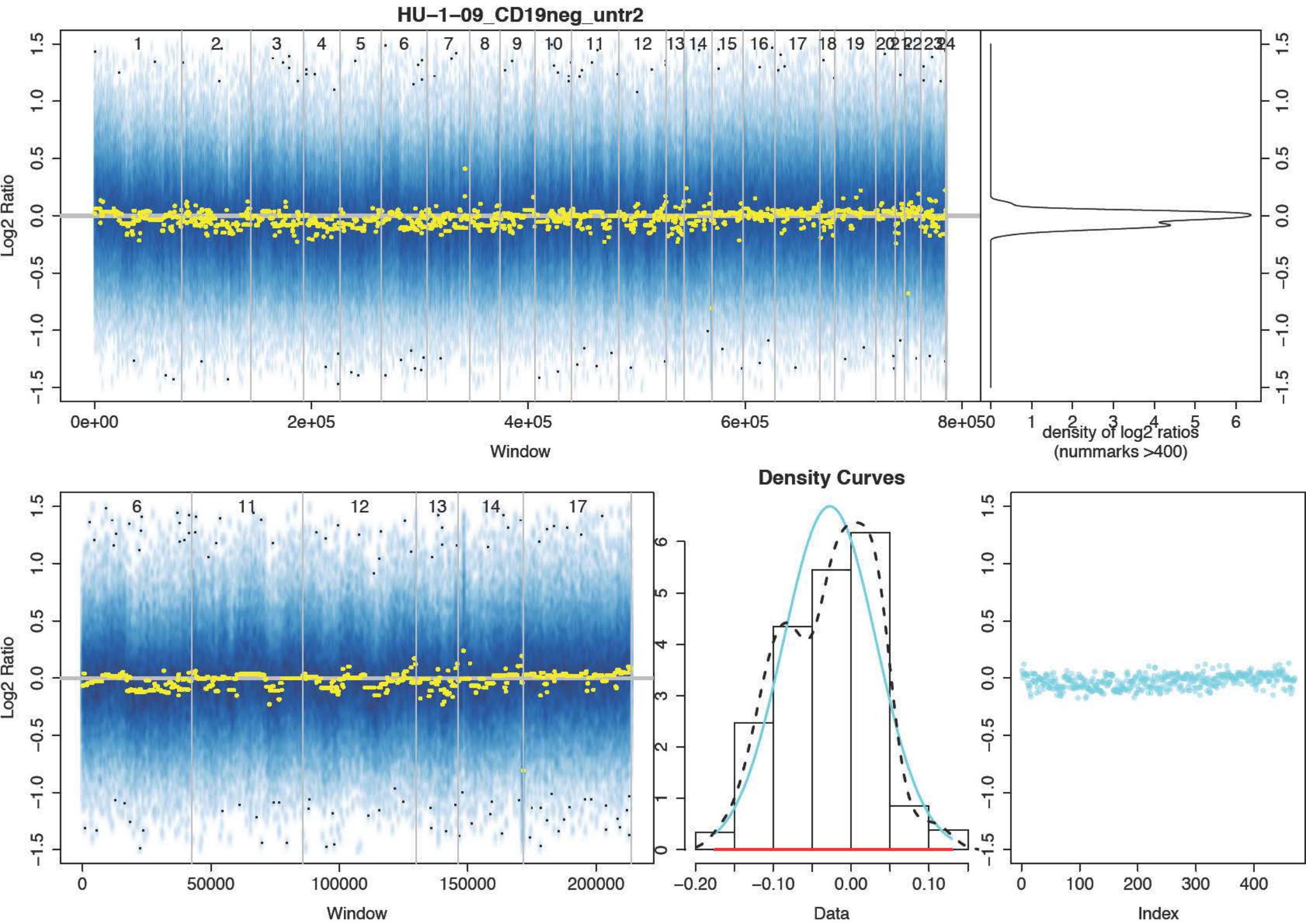

HU−1−07_normal6_rel
Log2 Ratio
Window
density of log2 ratios
(nummarks >400)
Density Curves
Data
Index

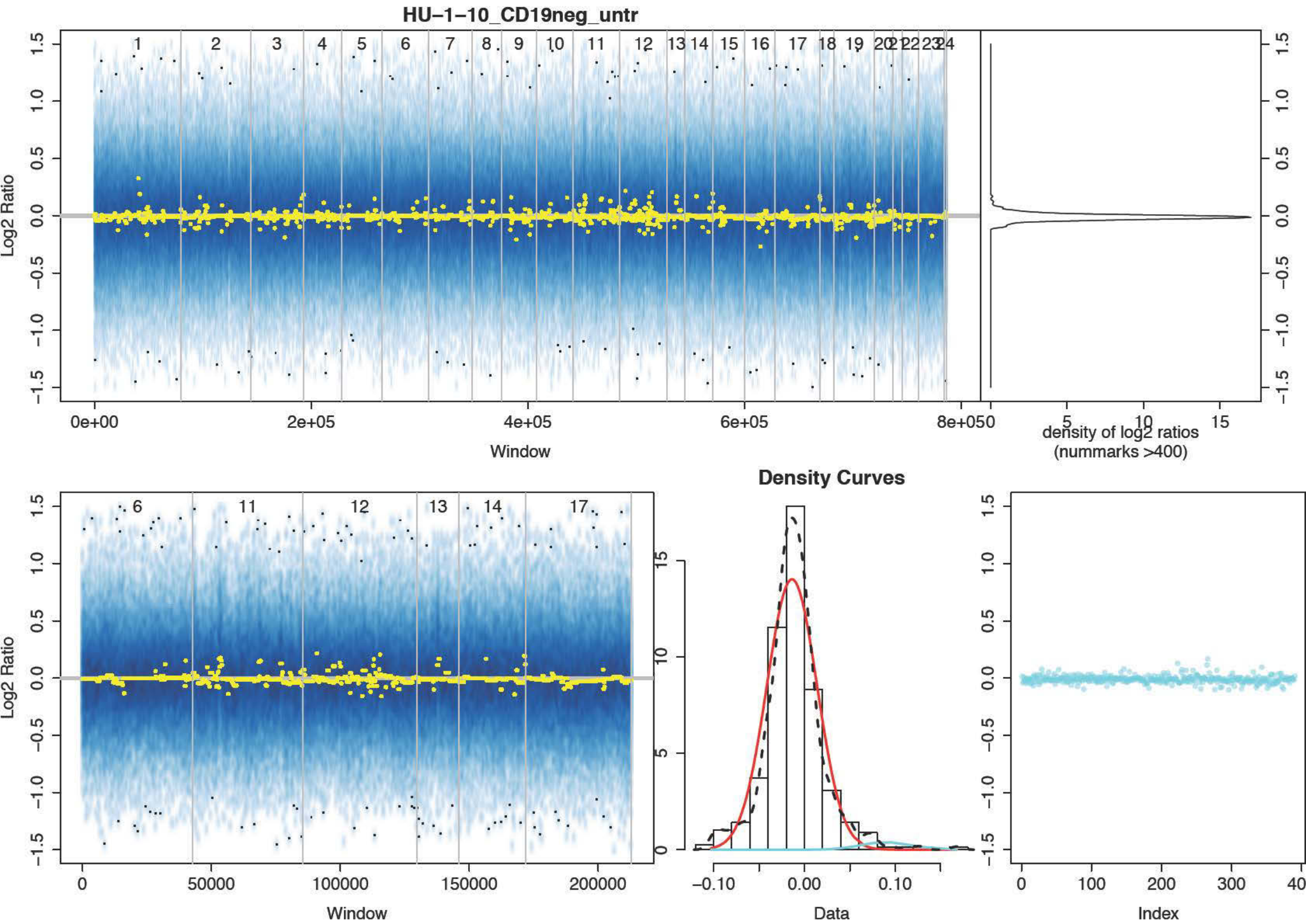

HU−1−08_normal6_untr
Log2 Ratio
Window
density of log2 ratios
(nummarks >400)
Density Curves
Data
Index

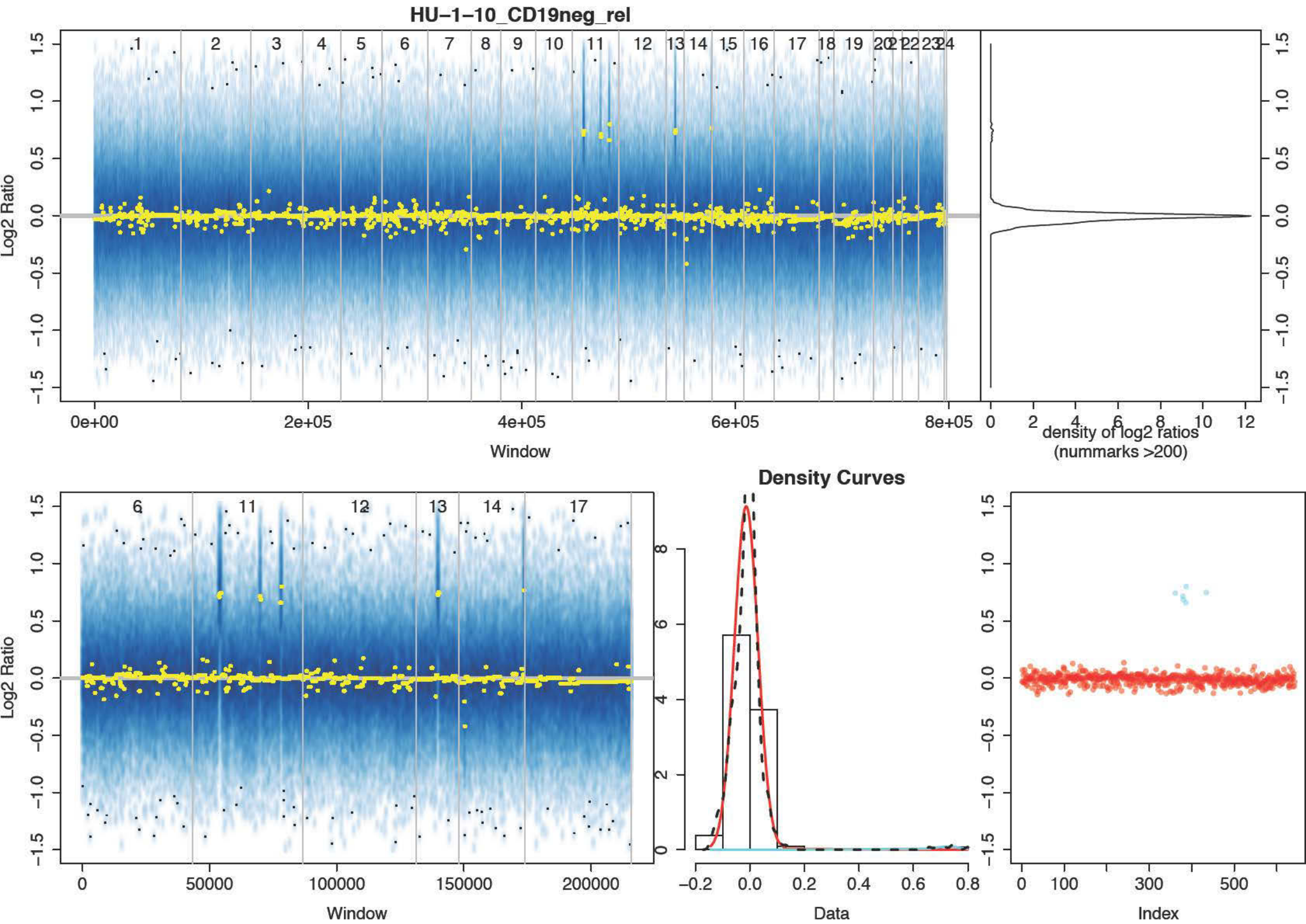

HU−1−08_CD19neg_untr2
Log2 Ratio
Window
0e+00
2e+05
4e+05
6e+05
8e+05
Density Curve
Data
Index
Window
0
50000
150000
−1.0
−0.4
0.2
0
300
600

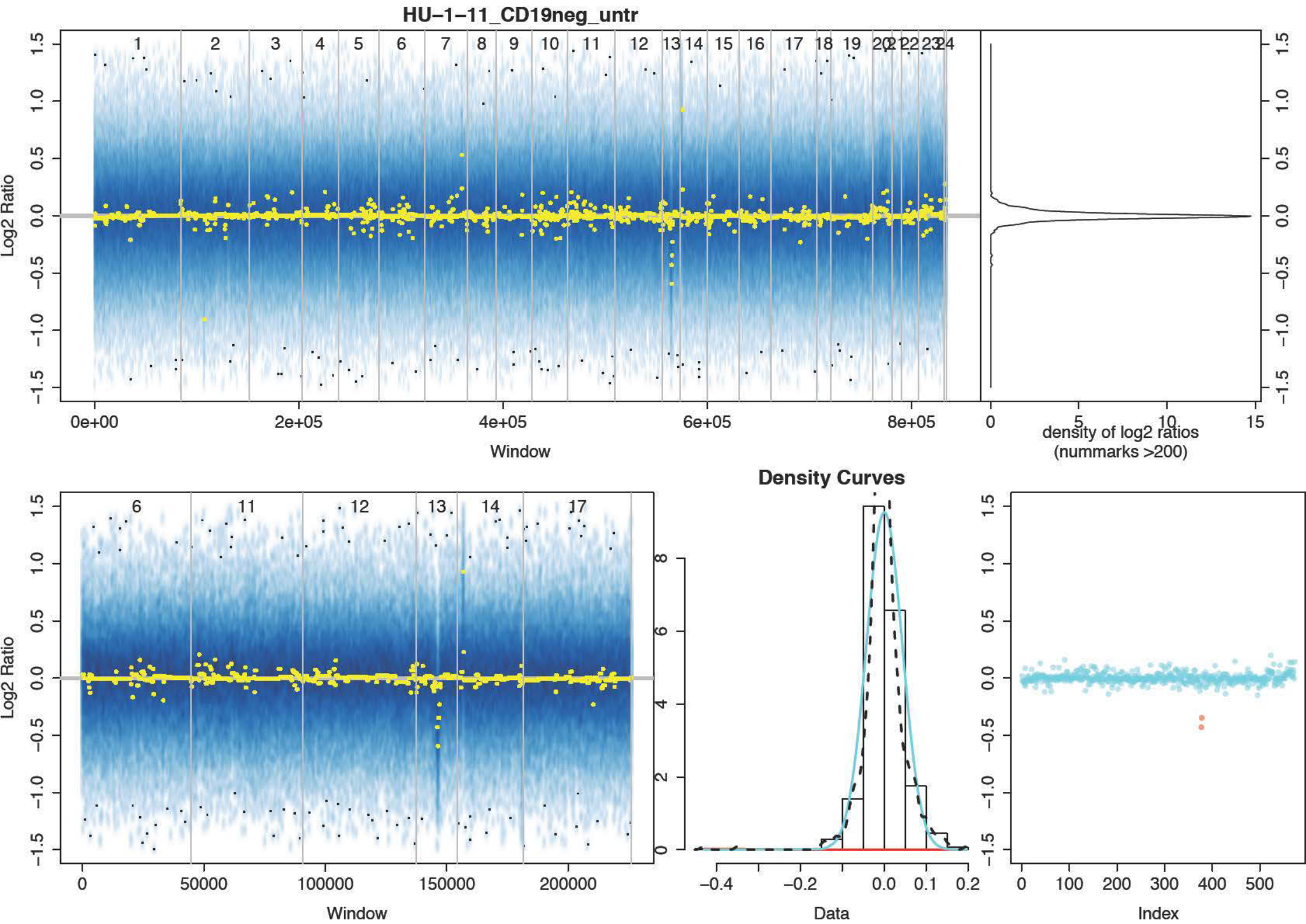

HU−1−09_CD19neg_untr
Log2 Ratio
Window
0e+00
2e+05
4e+05
6e+05
8e+05
sity of log2 ratios
ummarks >400)

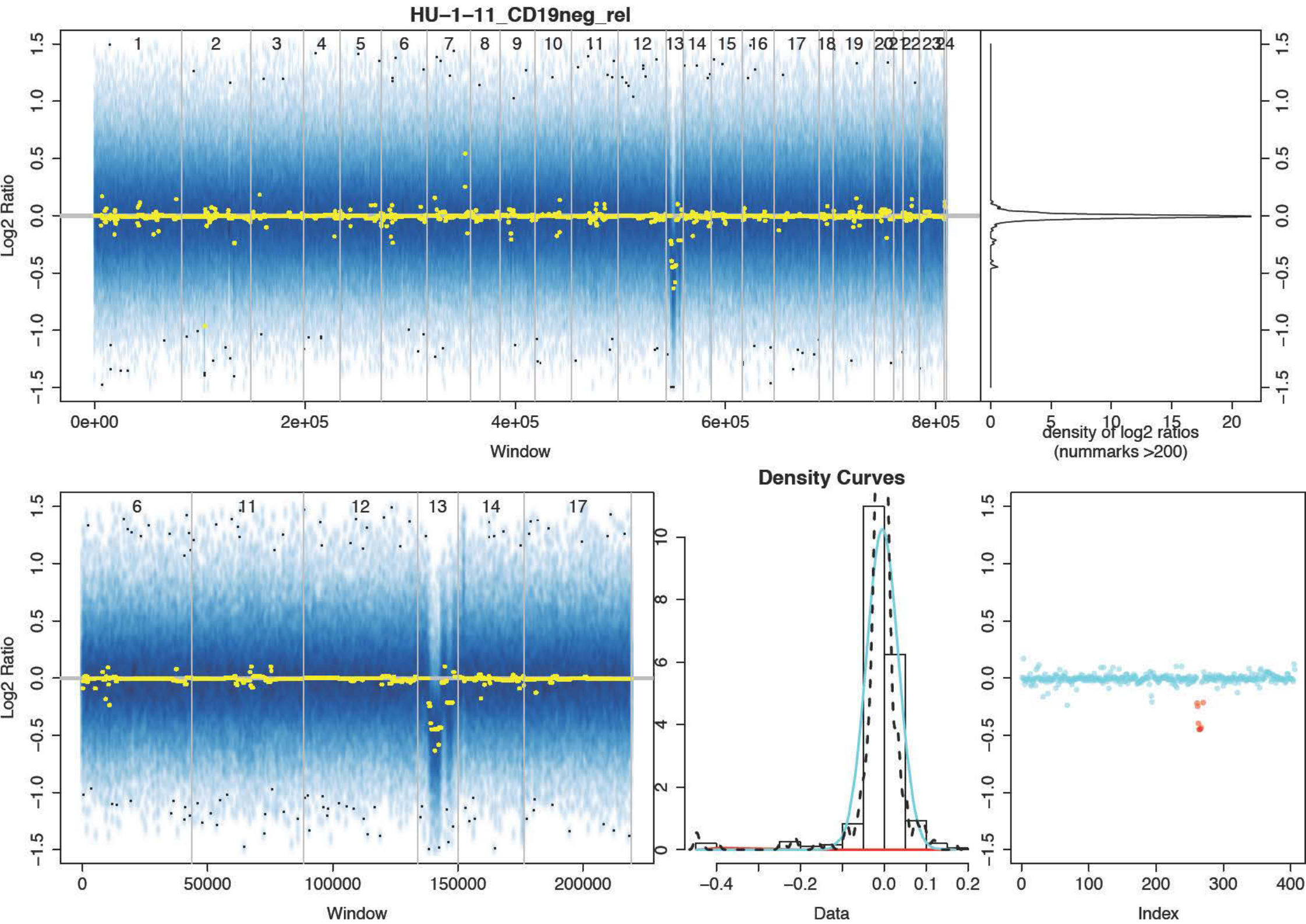

Density Curve
Log2 Ratio
Window
Data
Index

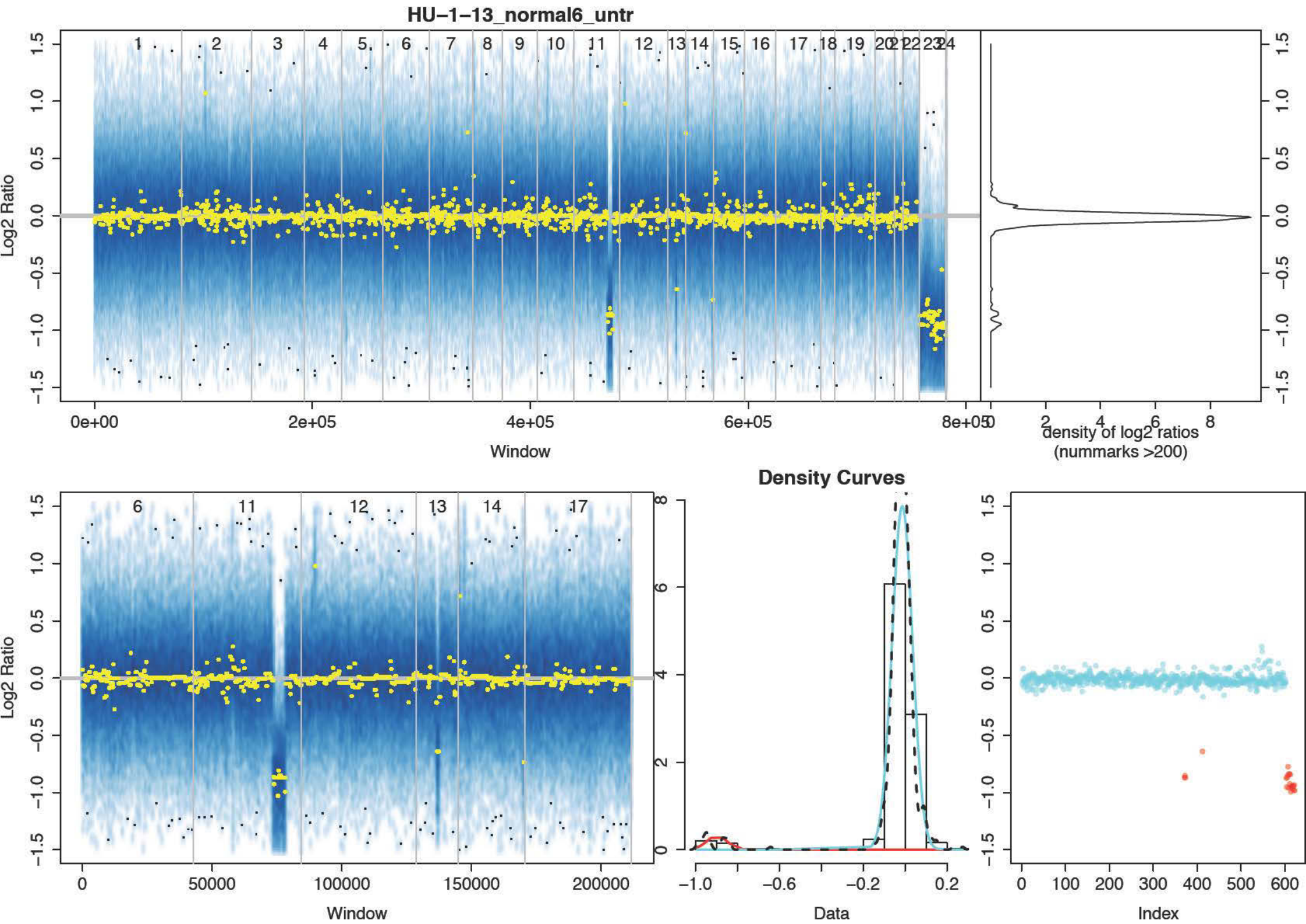

HU−1−09_CD19neg_untr2
Log2 Ratio
Window
density of log2 ratios
(nummarks >400)
Density Curves
Data
Index

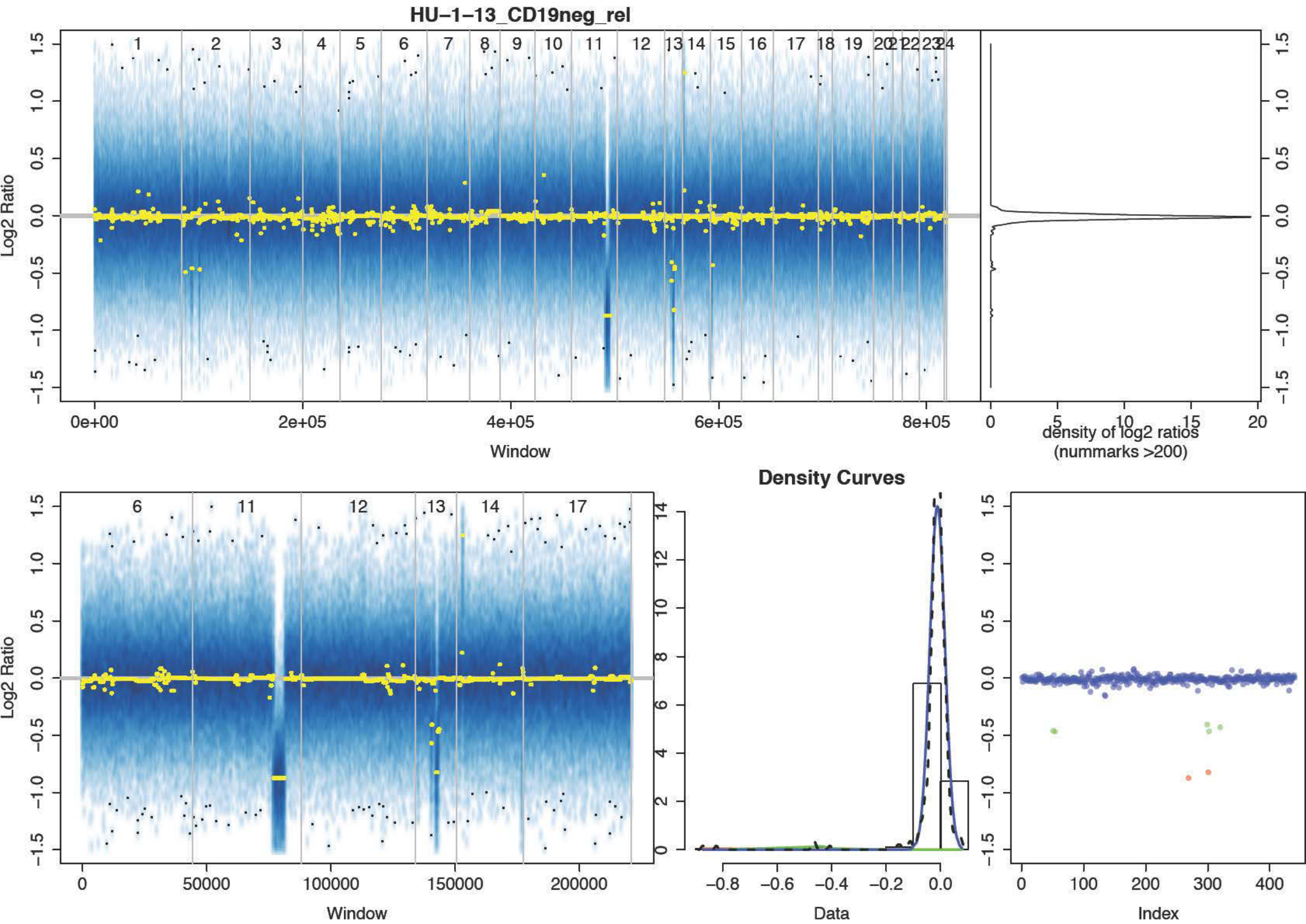

HU−1−10_CD19neg_untr
Log2 Ratio
Window
density of log2 ratios
(nummarks >400)
Density Curves
Data
Index

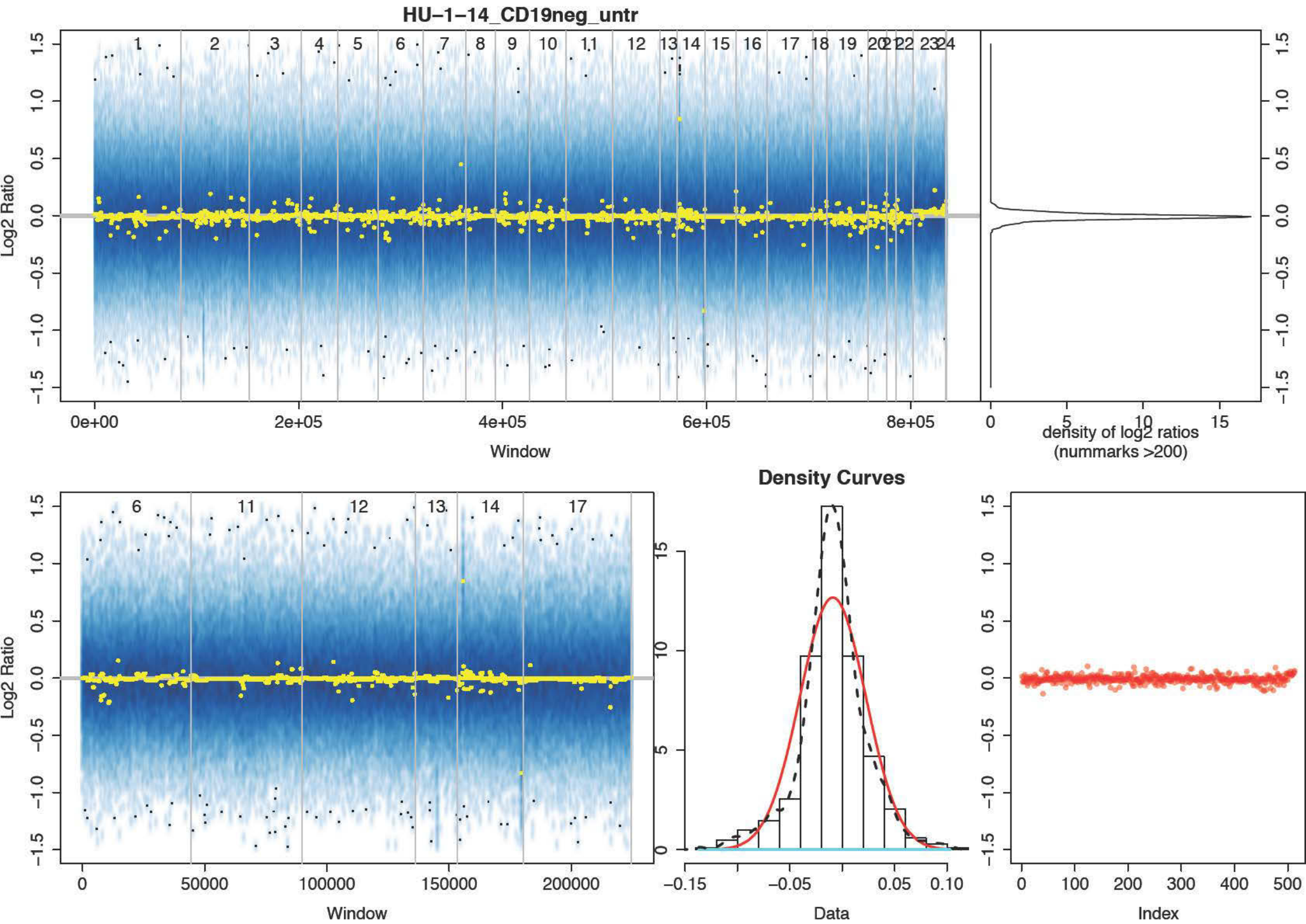

HU−1−10_CD19neg_rel
Log2 Ratio
Window
density of log2 ratios
(nummarks >200)
Density Curves
Data
Index

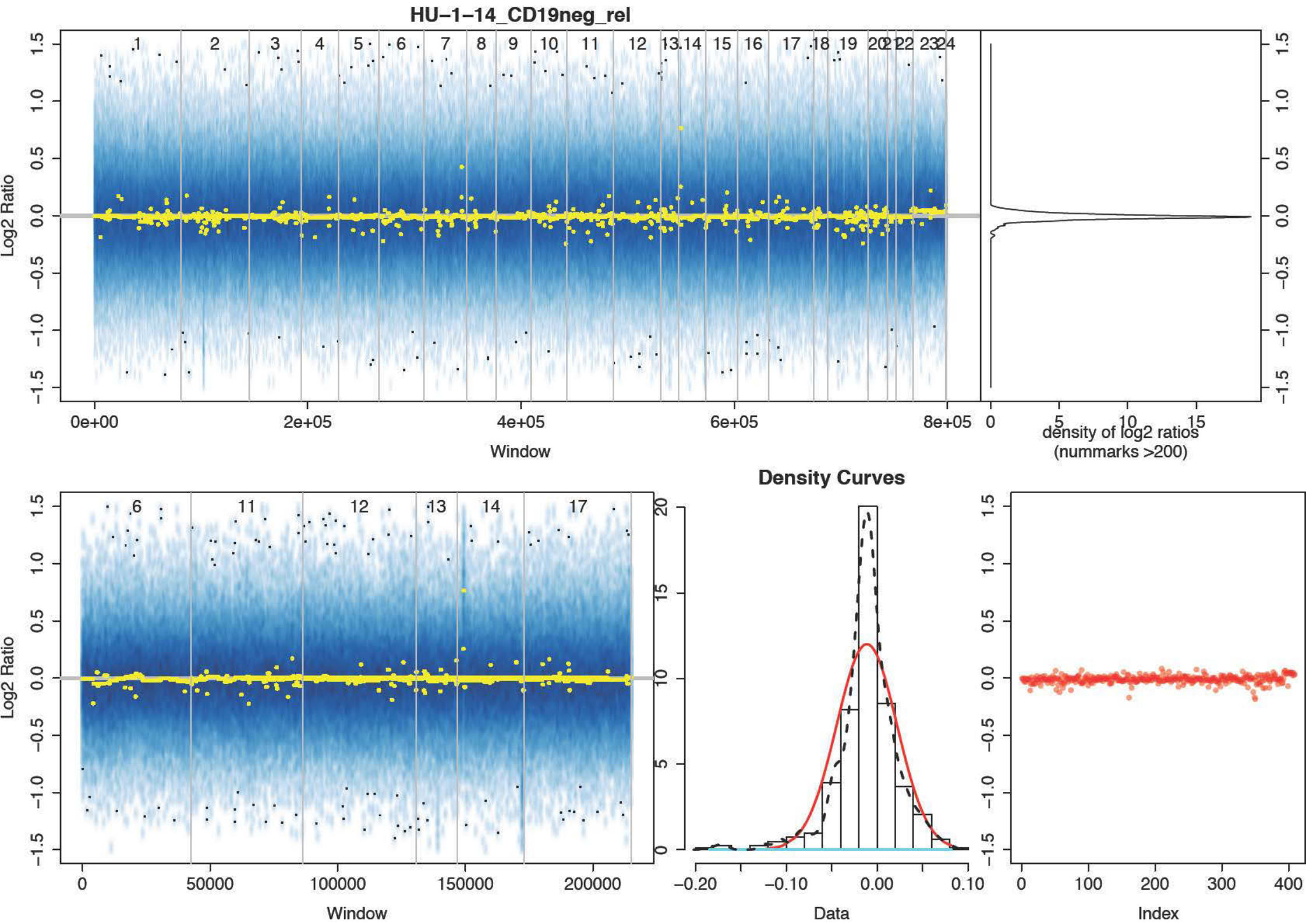

HU−1−11_CD19neg_untr
Log2 Ratio
Window
density of log2 ratios
(nummarks >200)
Density Curves
Data
Index

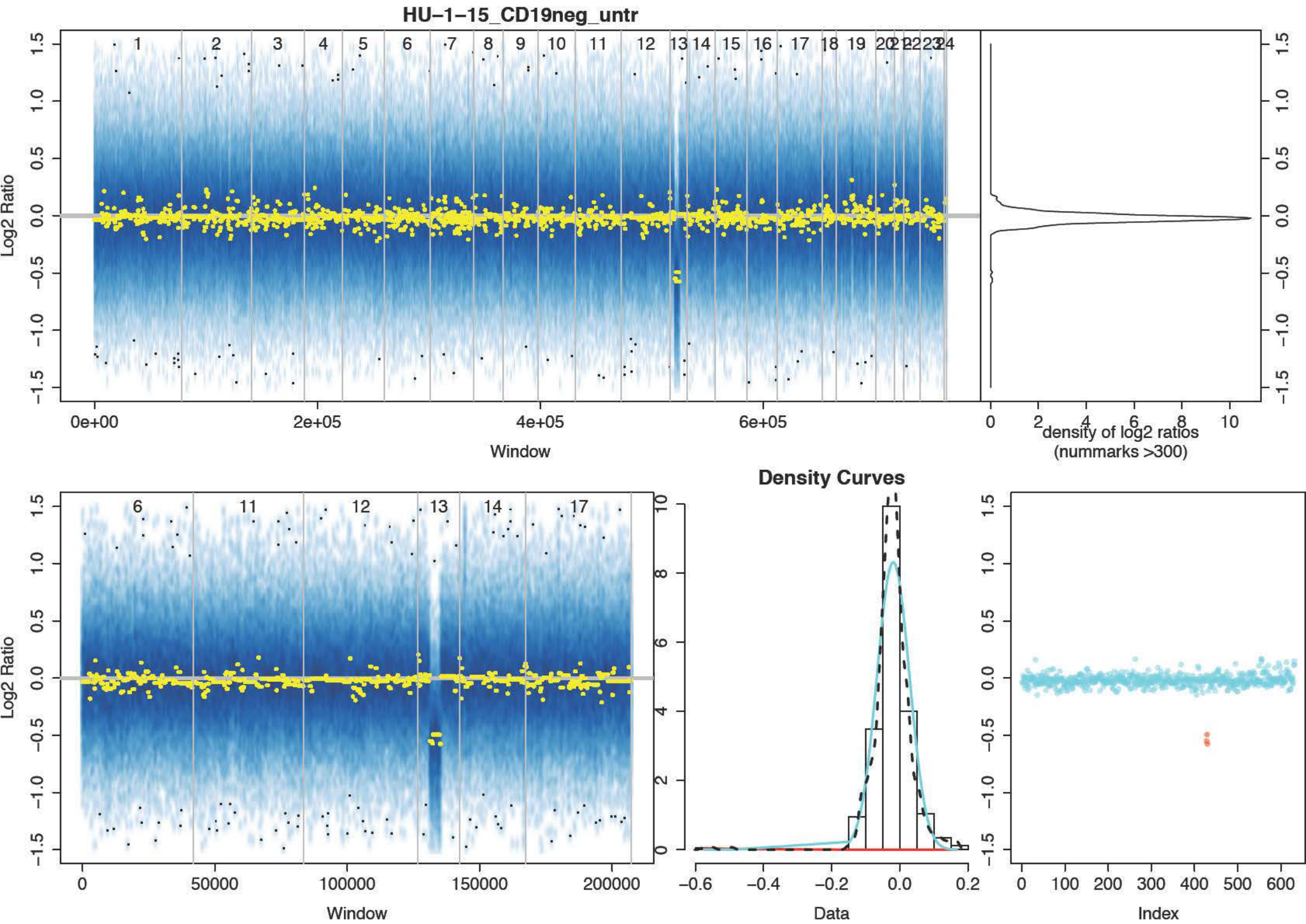

HU−1−11_CD19neg_rel
Log2 Ratio
Window
density of log2 ratios
(nummarks >200)
Density Curves
Data
Index

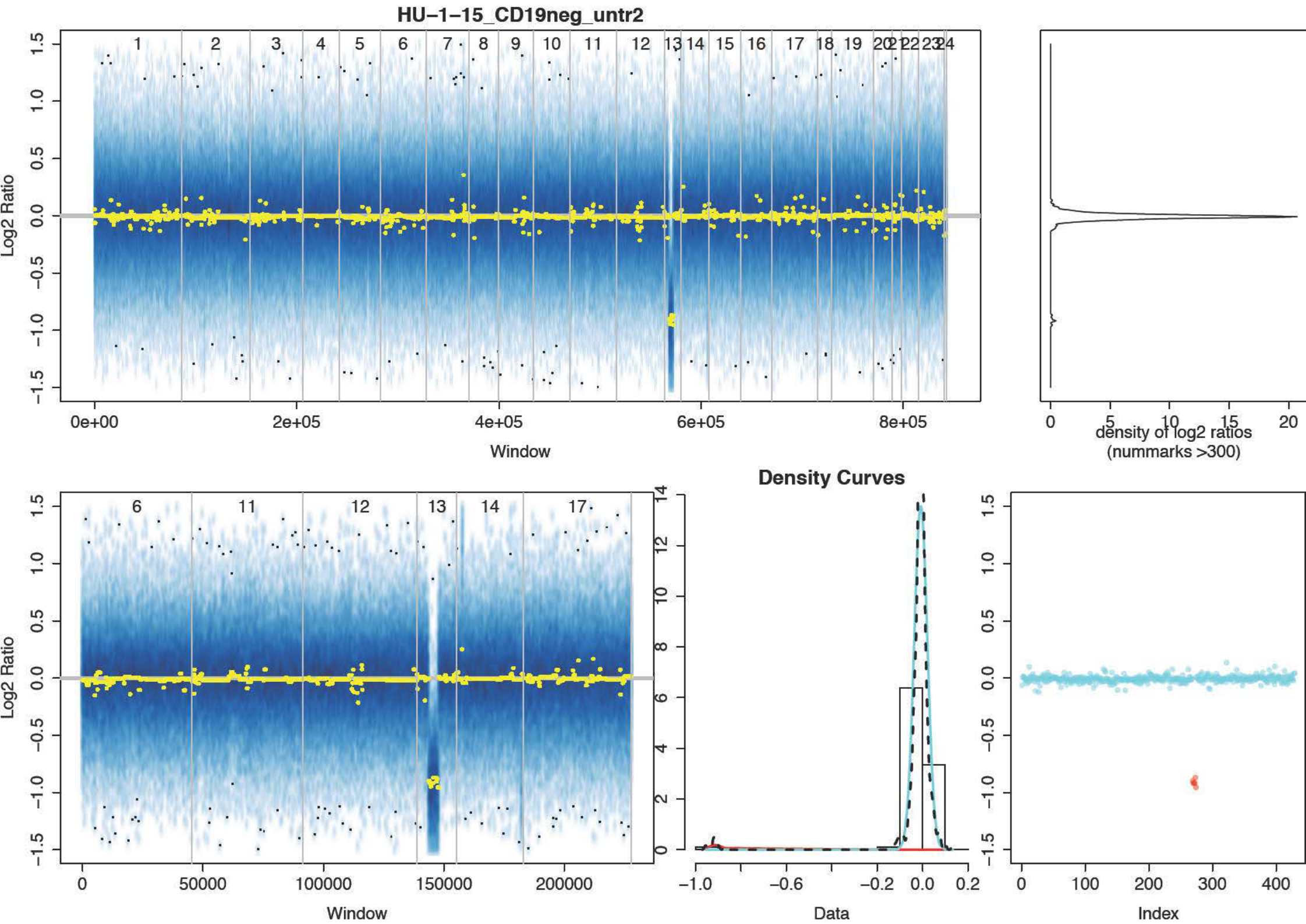

HU−1−13_normal6_untr
Log2 Ratio
Window
density of log2 ratios
(nummarks >200)
Density Curves
Data
Index

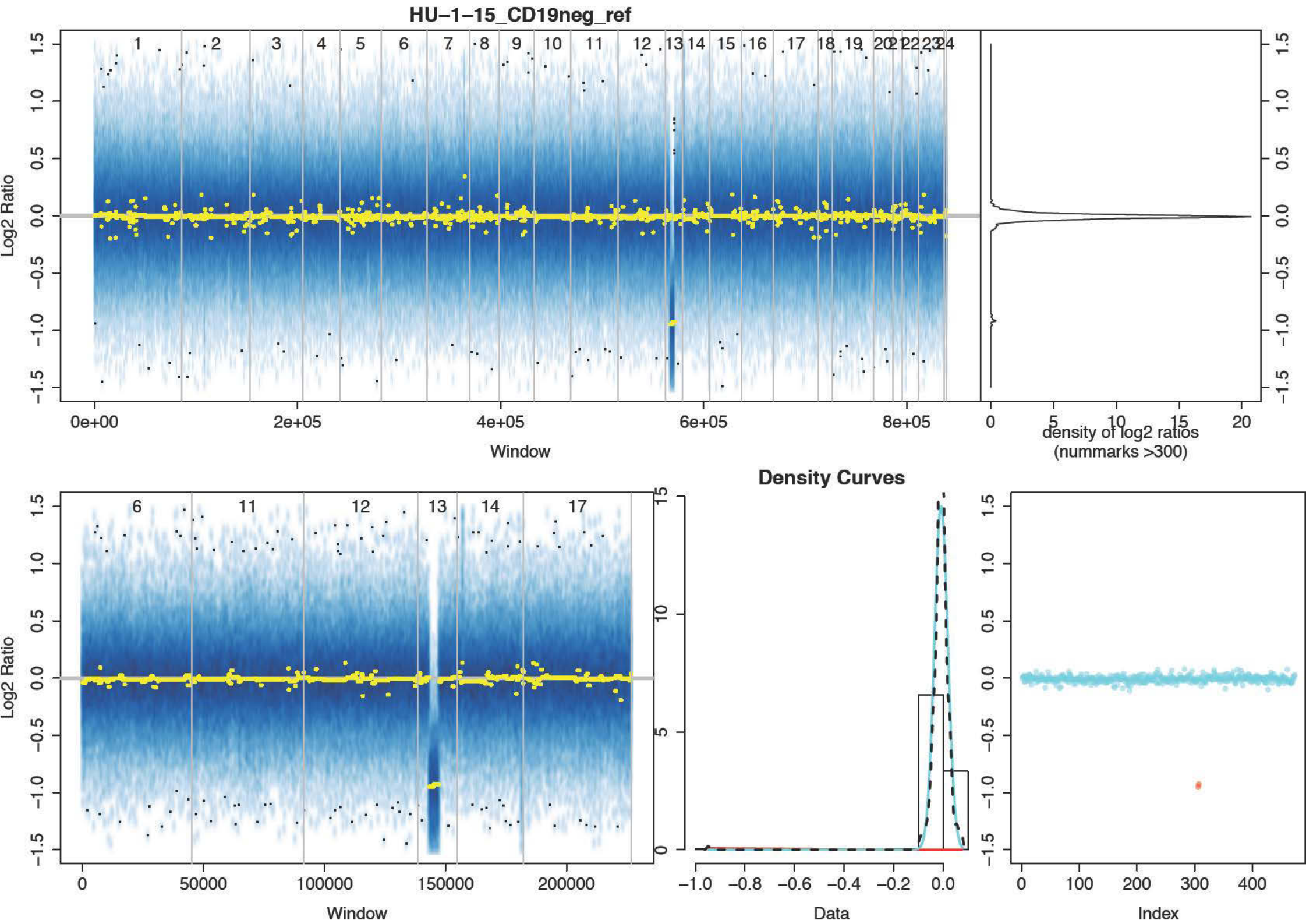

HU−1−13_CD19neg_rel
Log2 Ratio
Window
density of log2 ratios
(nummarks >200)
Density Curves
Data
Index

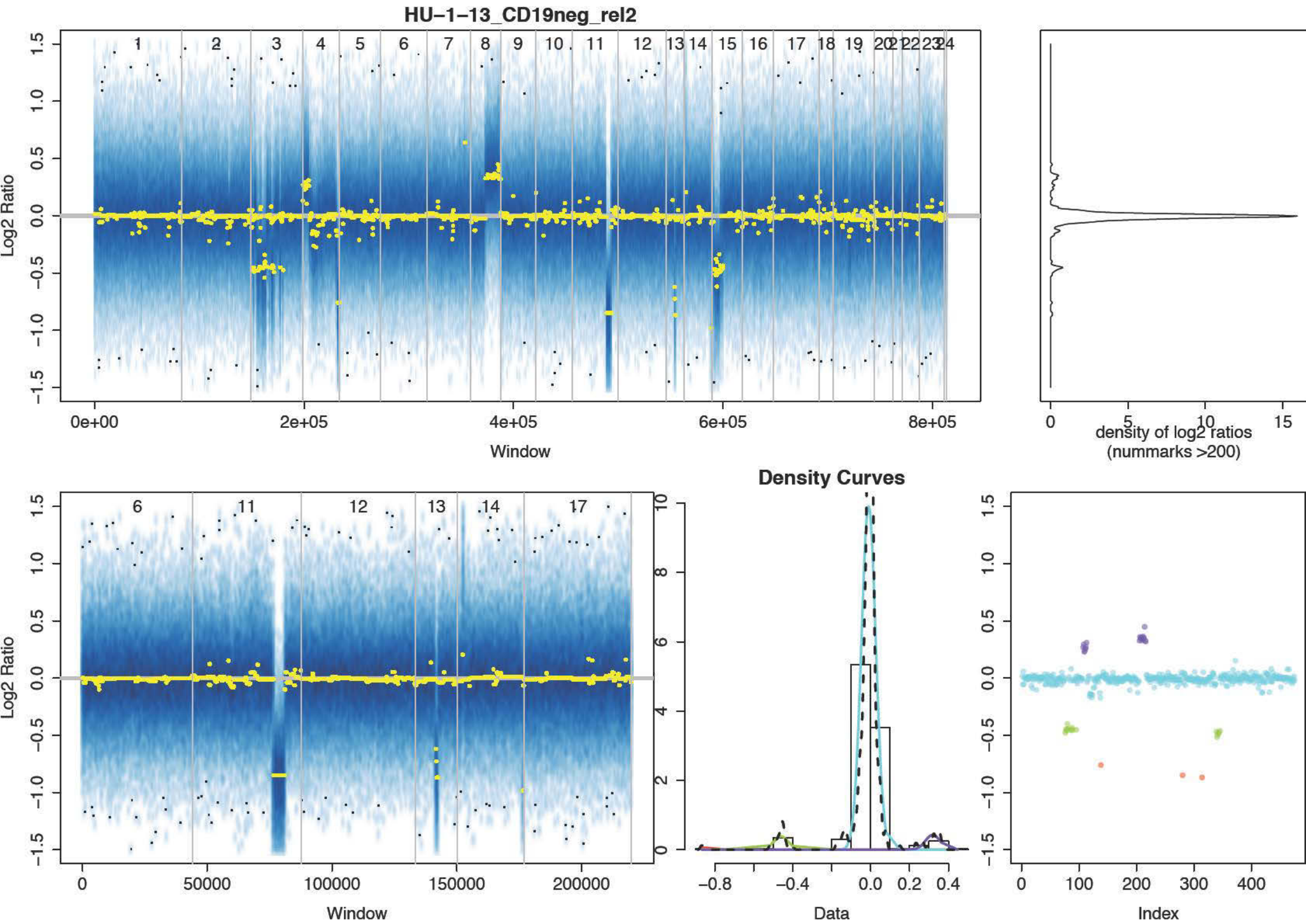

HU−1−14_CD19neg_untr
Log2 Ratio
Window
density of log2 ratios
(nummarks >200)
Density Curves
Data
Index

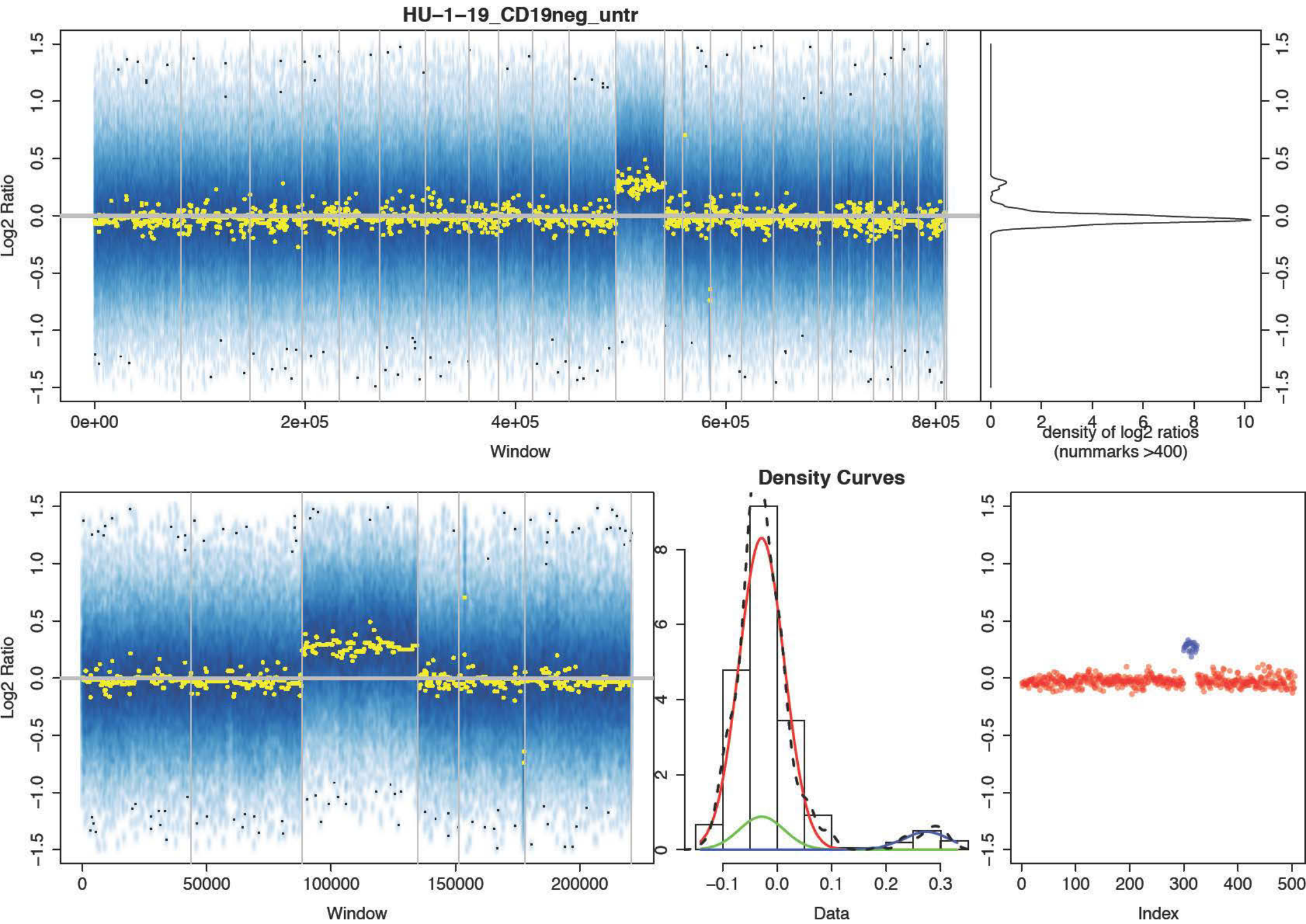

HU−1−14_CD19neg_rel
Log2 Ratio
Window
density of log2 ratios
(nummarks >200)
Density Curves
Data
Index

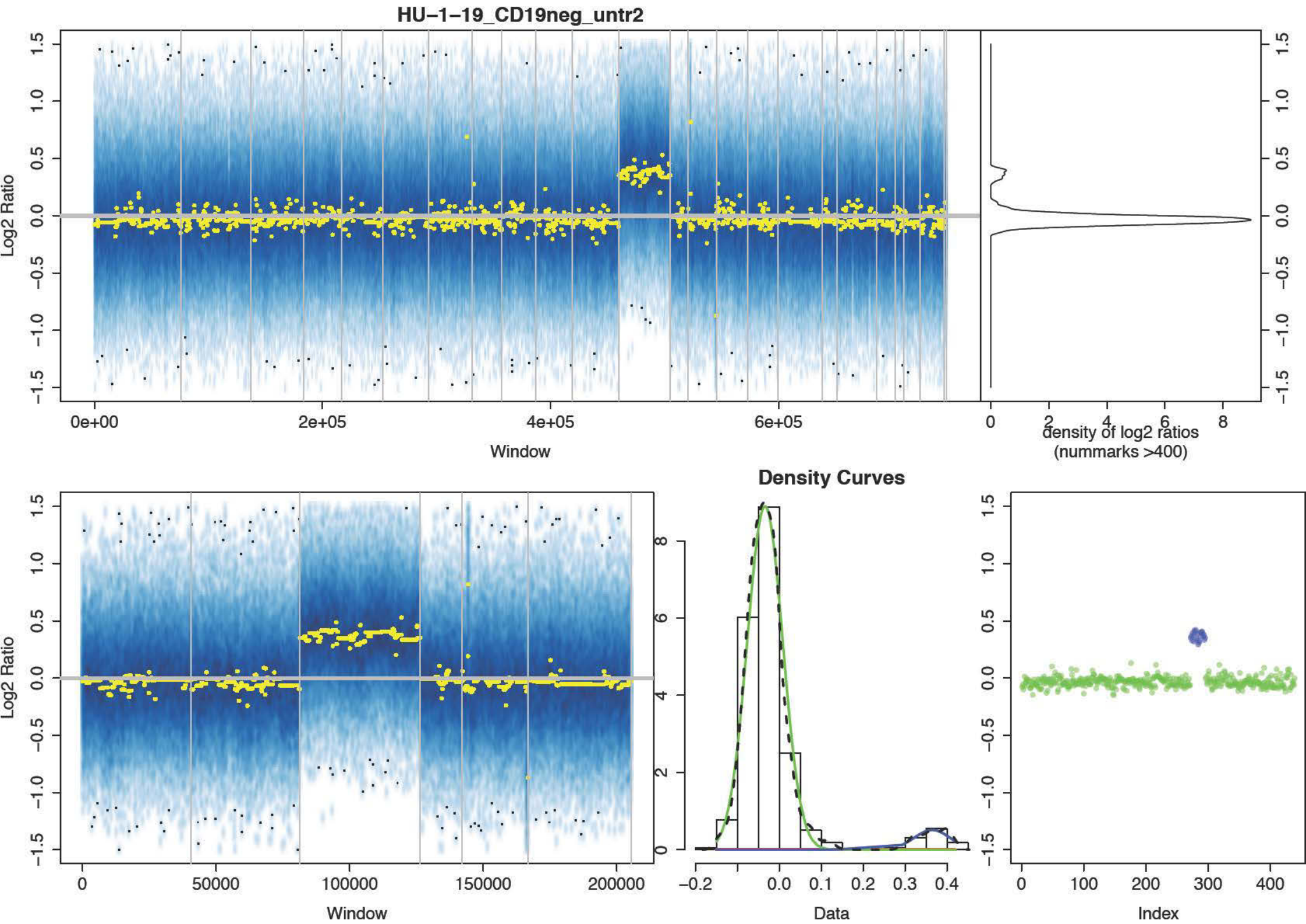

HU−1−15_CD19neg_untr
Log2 Ratio
Window
density of log2 ratios
(nummarks >300)
Density Curves
Data
Index

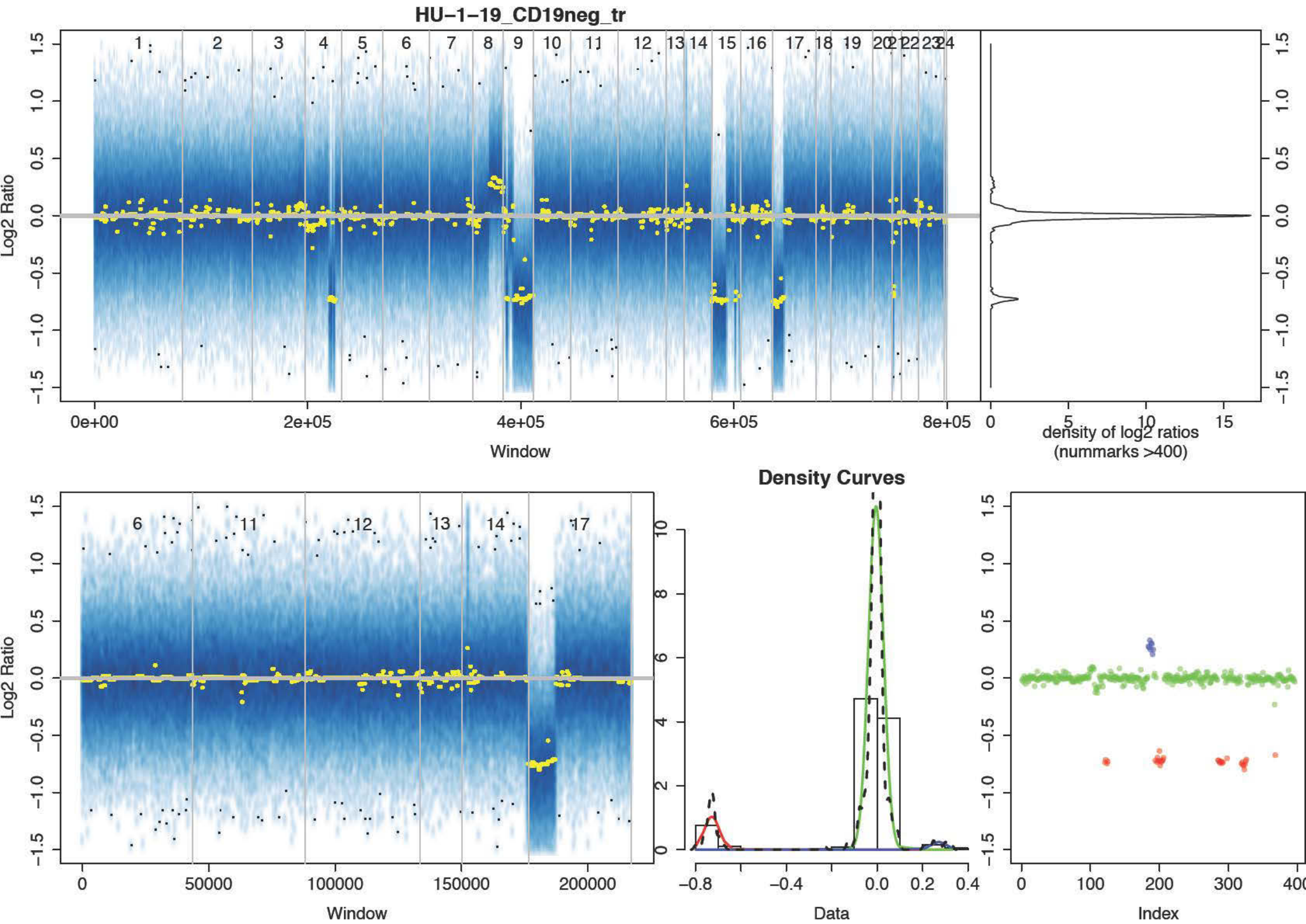

HU−1−15_CD19neg_untr2
Log2 Ratio
Window
density of log2 ratios
(nummarks >300)
Density Curves
Data
Index

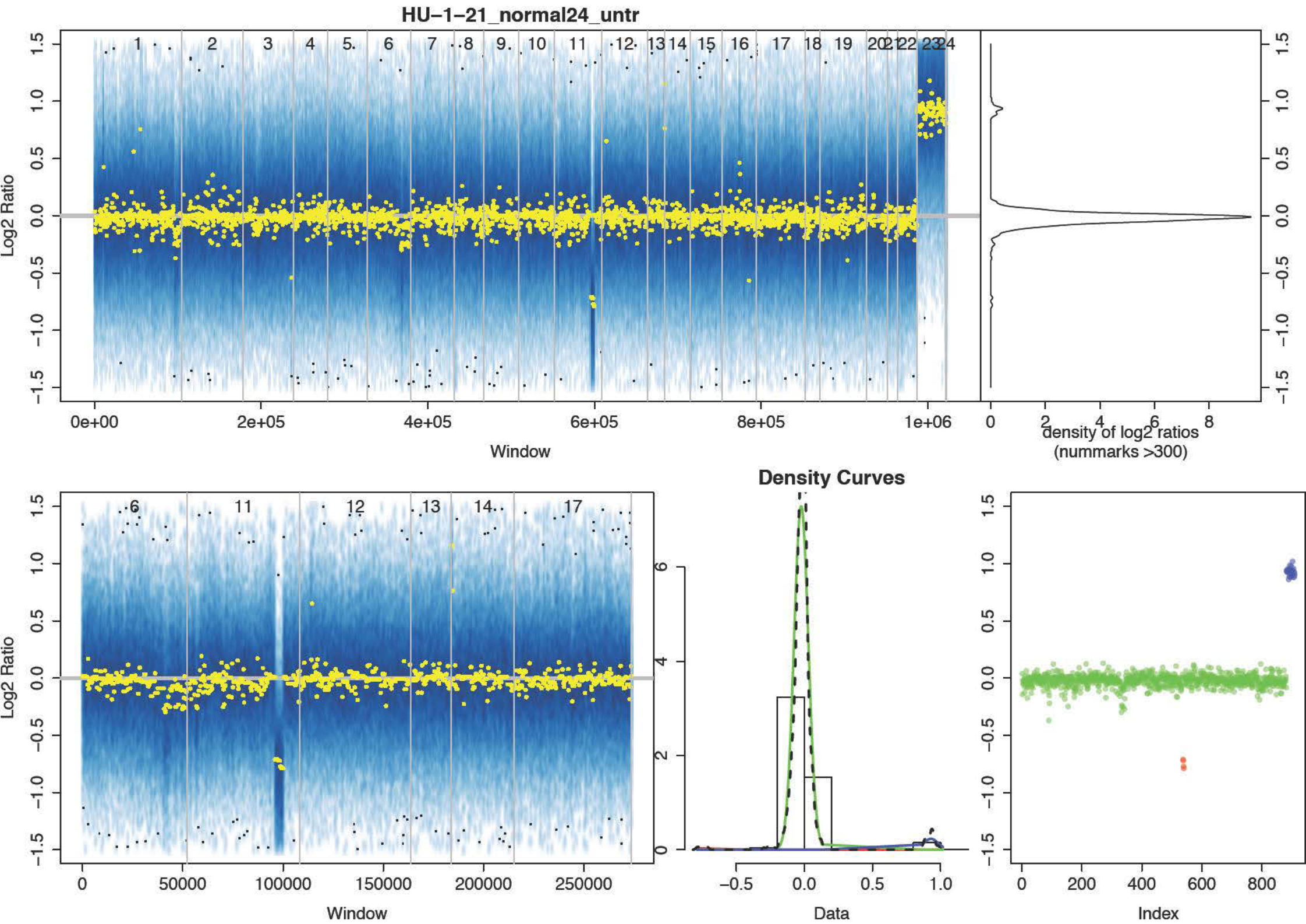

HU−1−15_CD19neg_ref
Log2 Ratio
Window
density of log2 ratios
(nummarks >300)
Density Curves
Data
Index

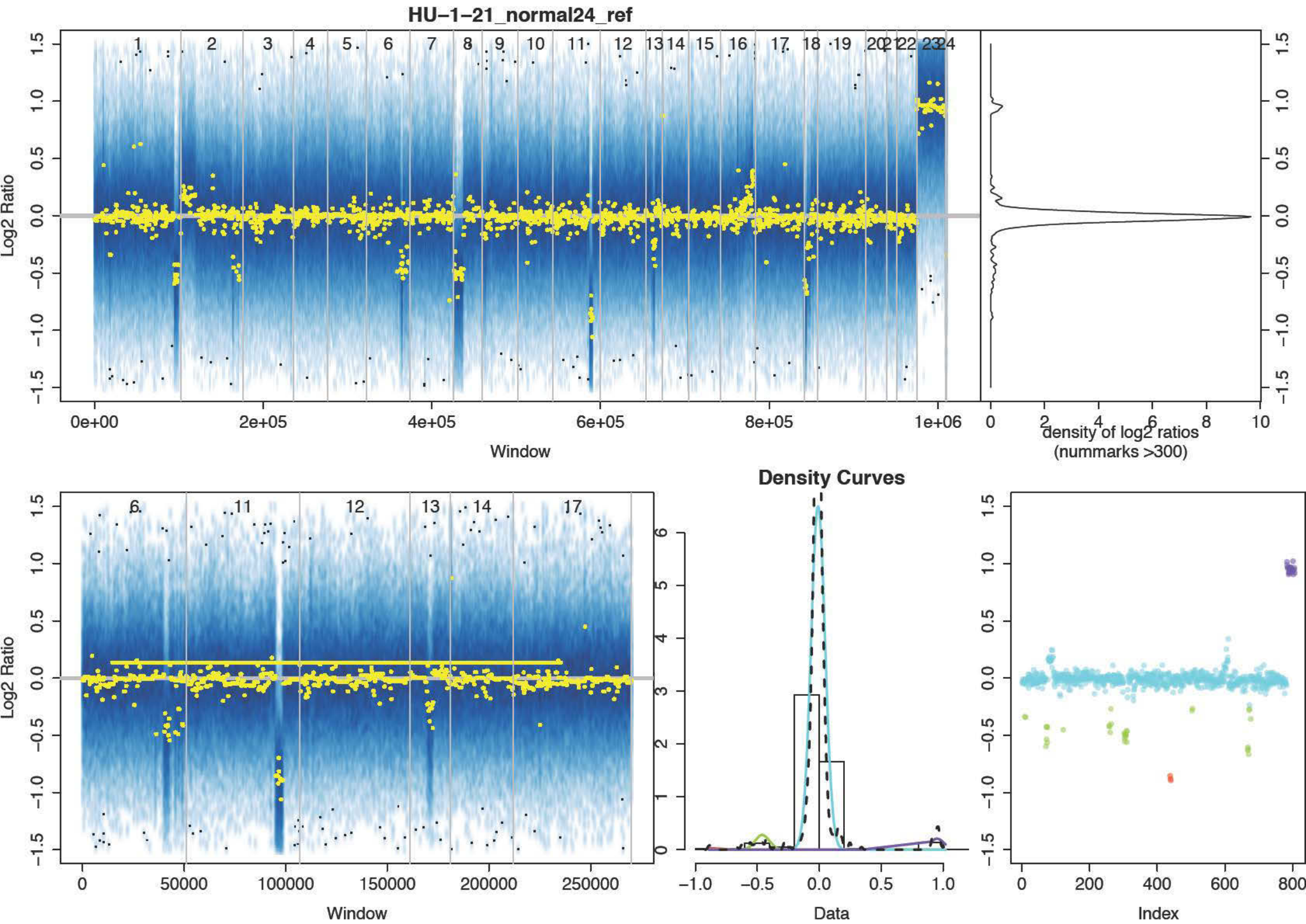

HU–1–13_CD19neg_rel2
Log2 Ratio
Window
density of log2 ratios
(nummarks >200)
Density Curves
Data
Index


HU−1−19_CD19neg_untr
Log2 Ratio
Window
density of log2 ratios
(nummarks >400)
Density Curves
Data
Index

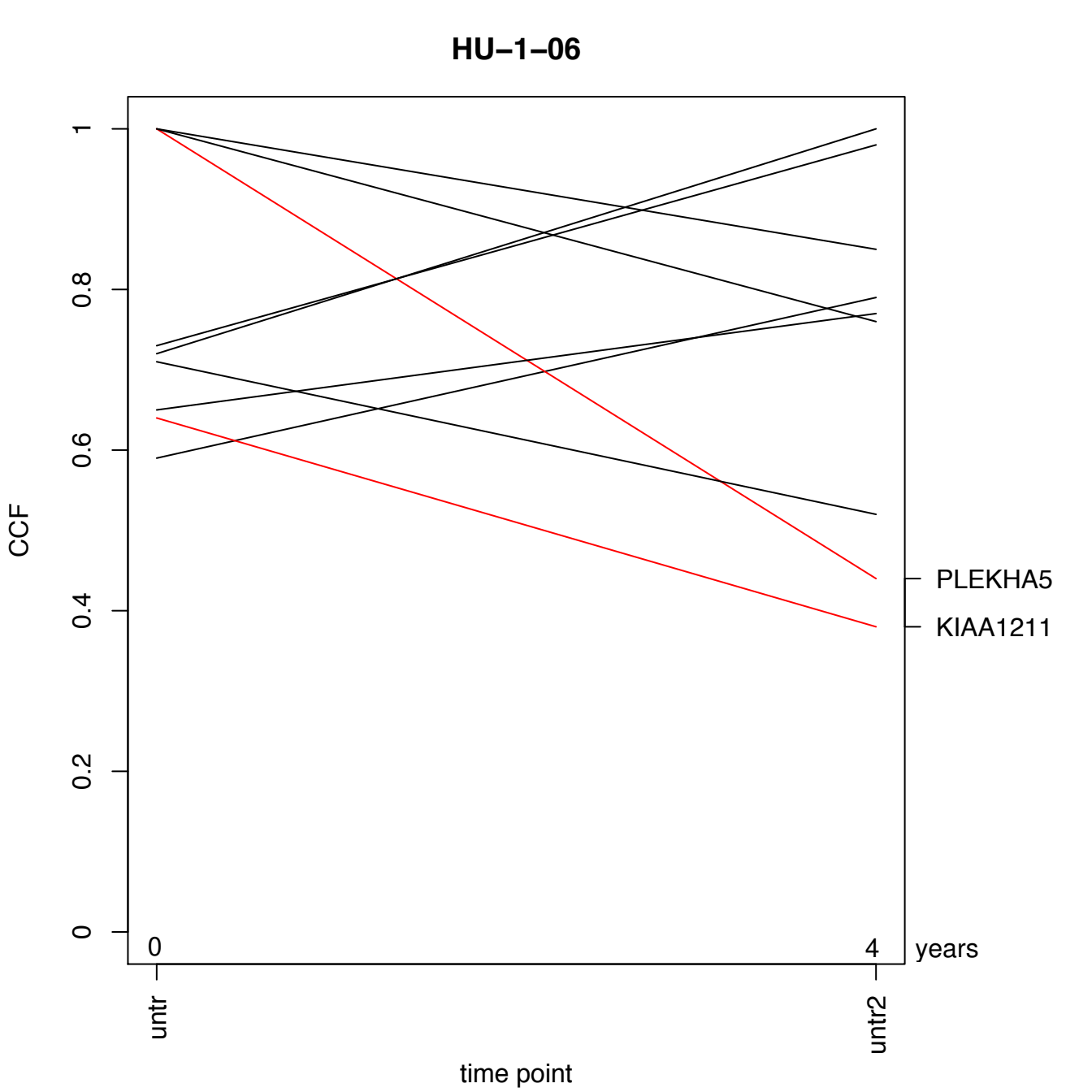

HU−1−19_CD19neg_untr2
Log2 Ratio
Window
density of log2 ratios
(nummarks >400)
Density Curves
Log2 Ratio
Window
Data
Index

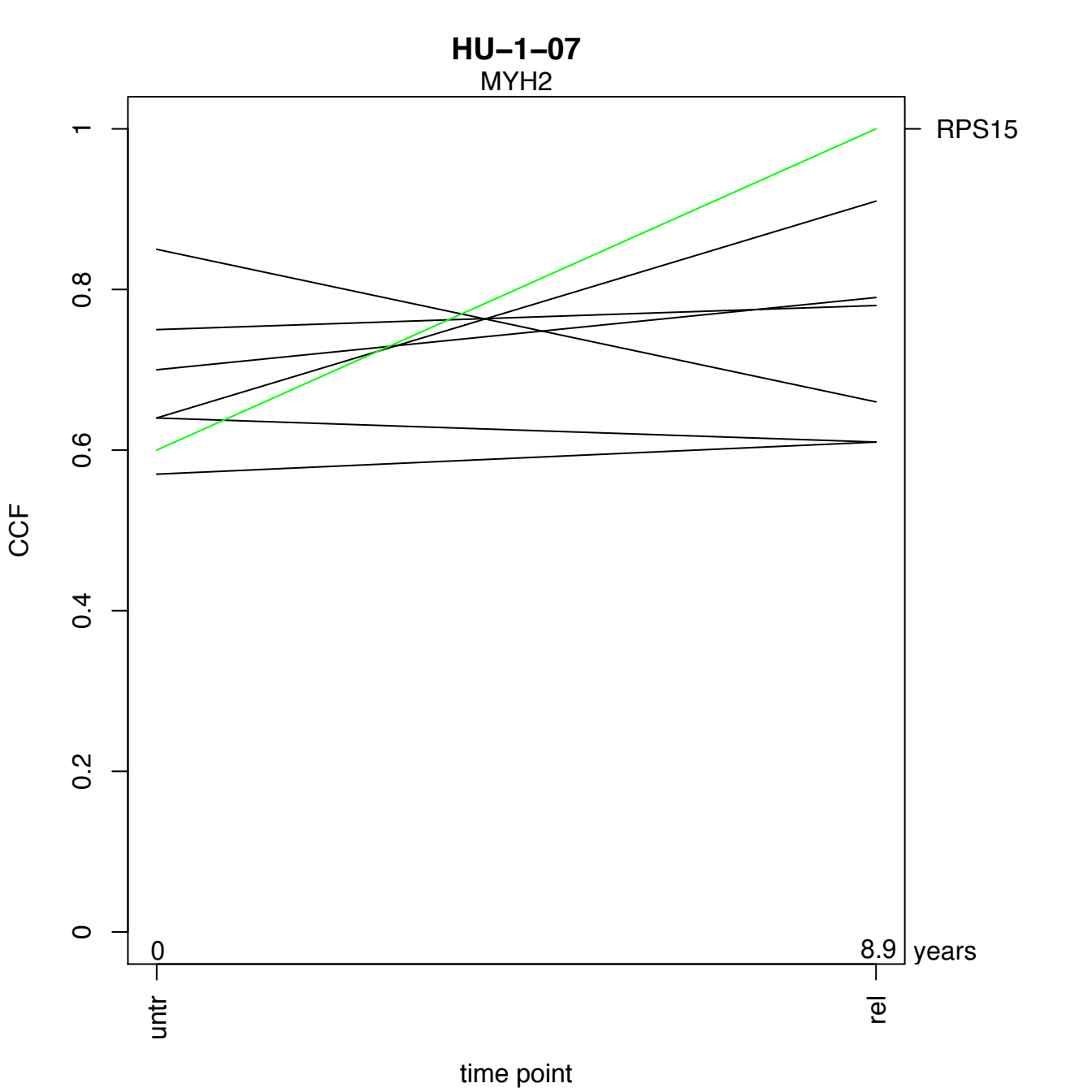

HU−1−19_CD19neg_tr
Log2 Ratio
Window
density of log2 ratios
(nummarks >400)
Density Curves
Data
Index

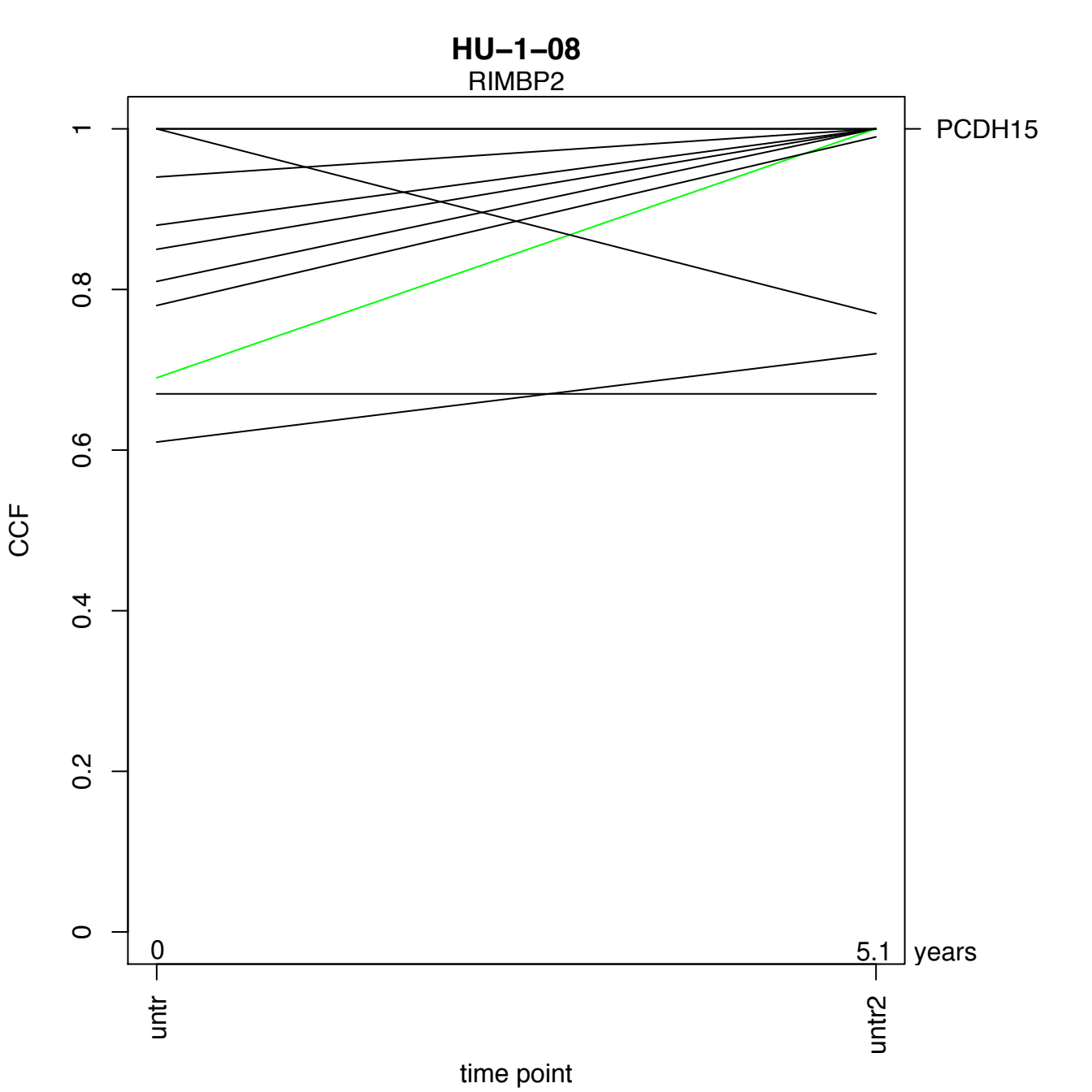

HU−1−21_normal24_untr
Log2 Ratio
Window
density of log2 ratios
(nummarks >300)
Density Curves
Data
Index

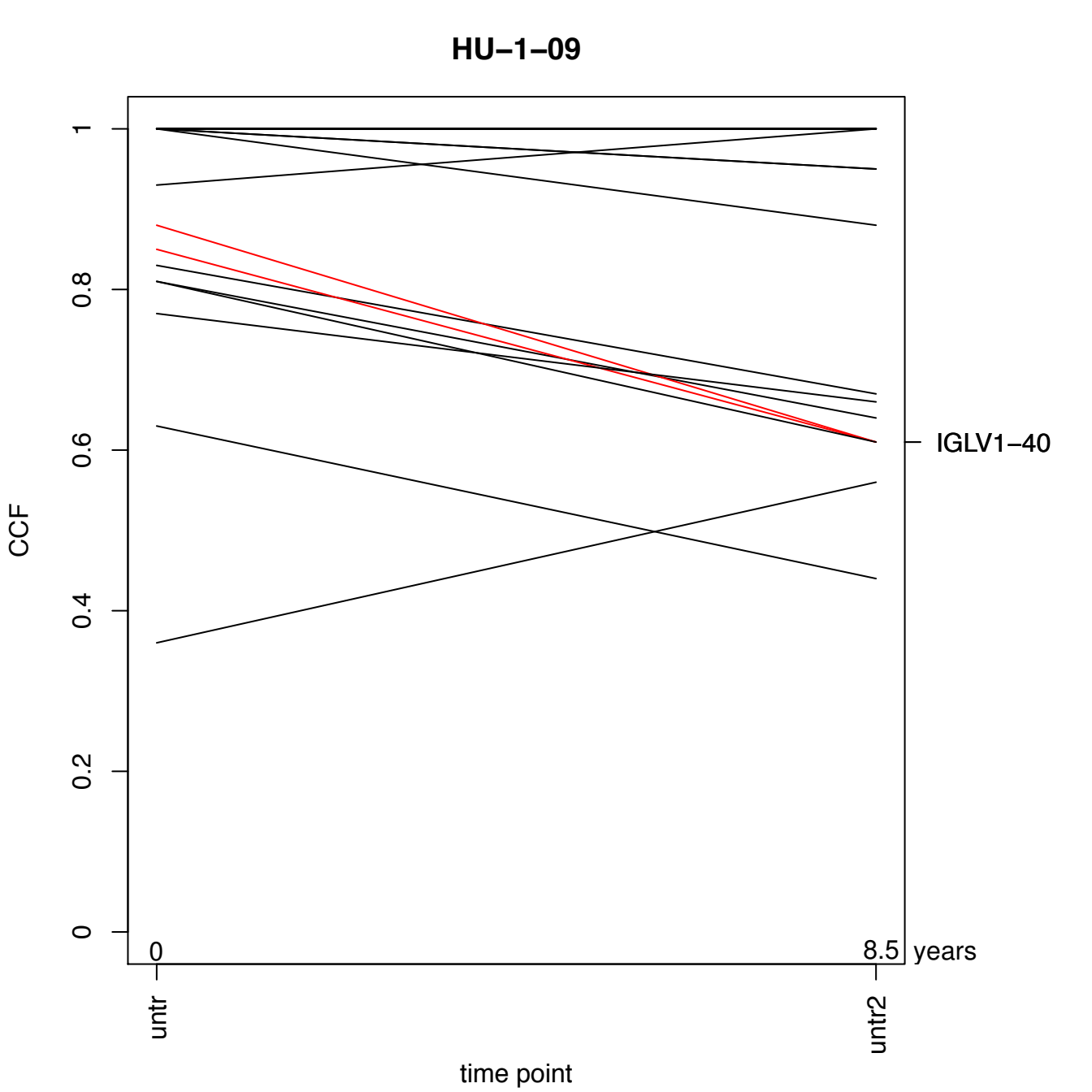

HU−1−21_normal24_ref
Log2 Ratio
Window
density of log2 ratios
(nummarks >300)
Density Curves
Data
Index

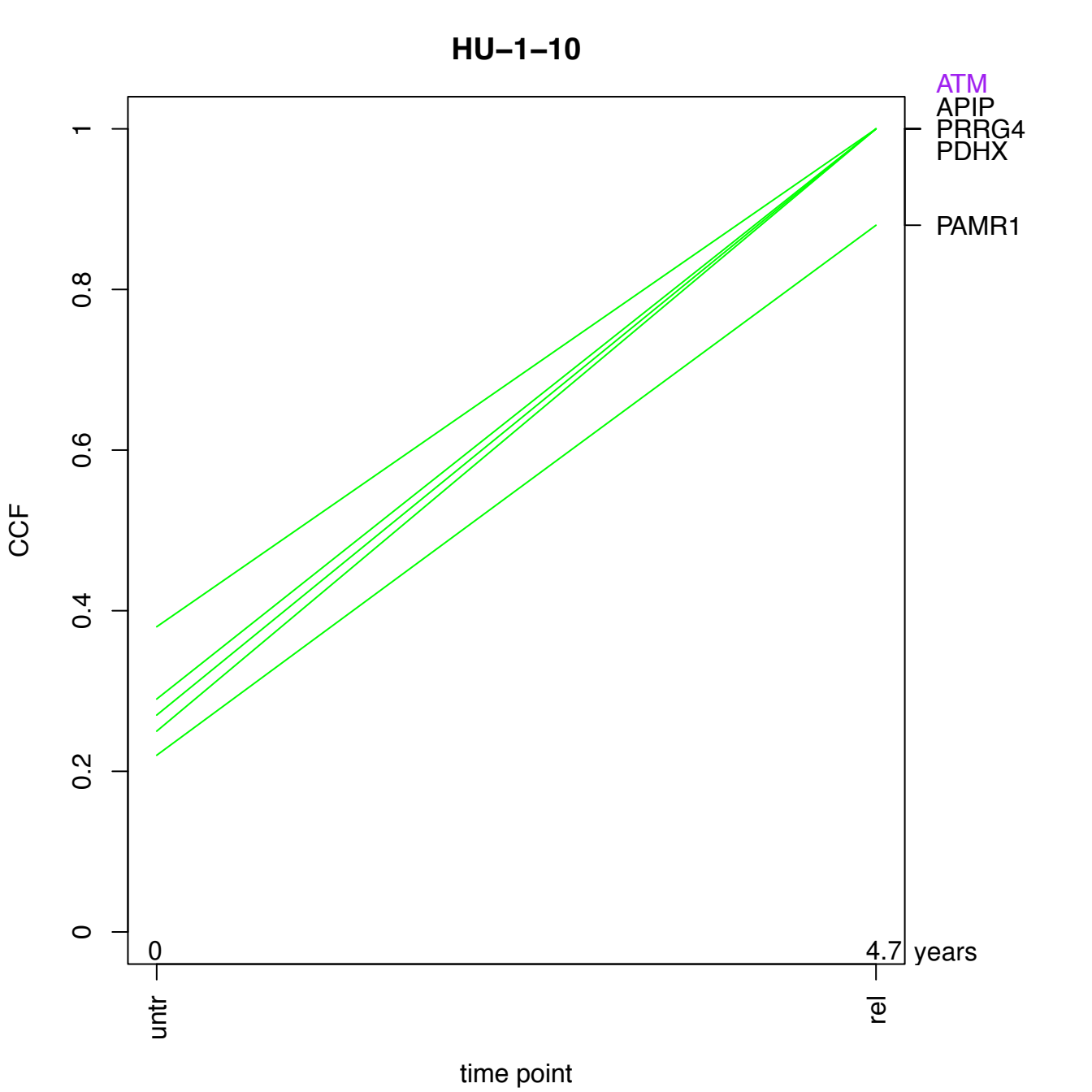
Suppl. figure 4: Changes in cancer cell fraction (CCF) with annotation of genes listed in the Catalogue of Somatic Mutations in Cancer (COSMIC). Significant changes in CCF are highlighted by red (reduced CCF) or green color (increased CCF). Genes represented in COSMIC are highlighted in purple.

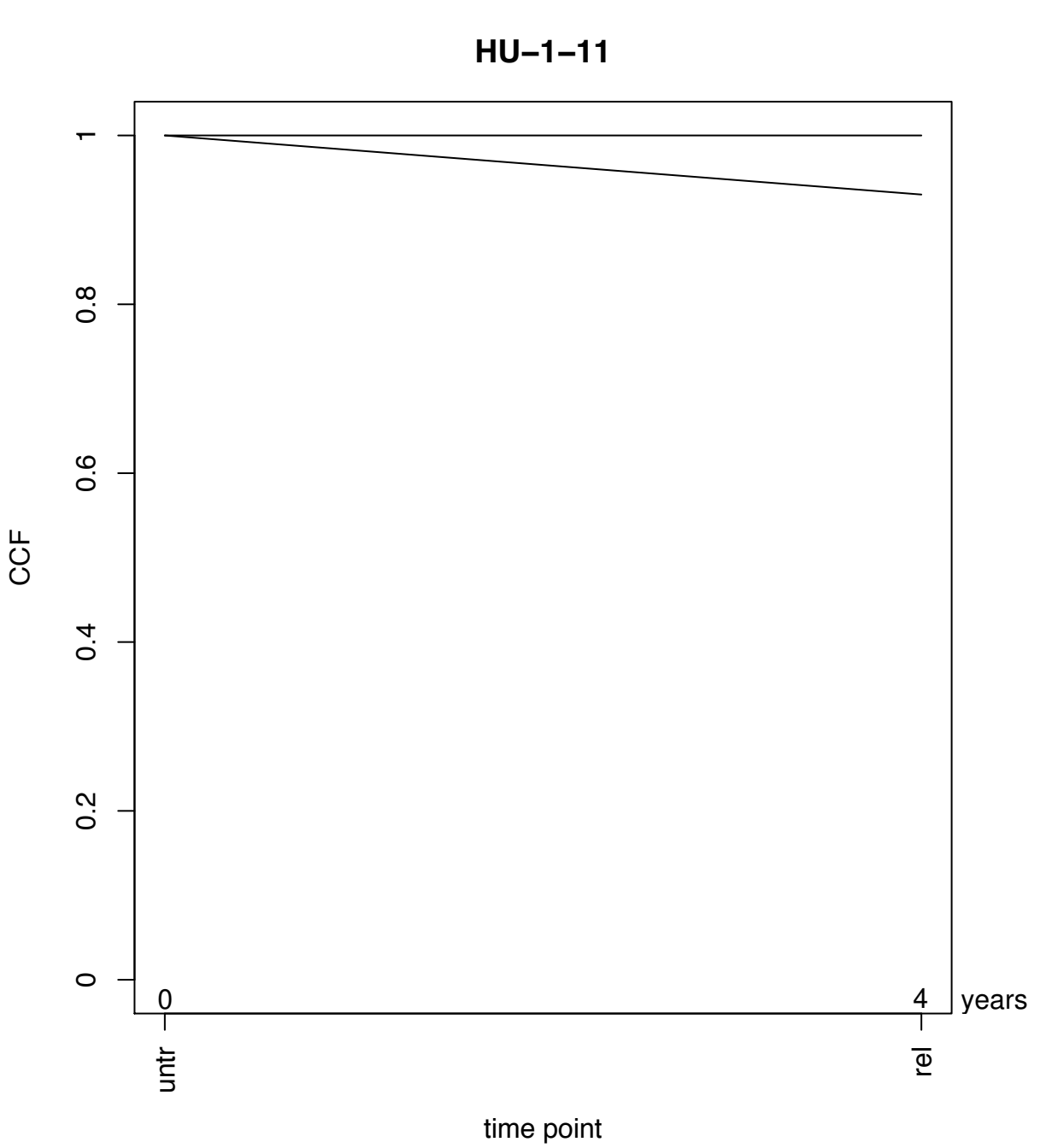

HU−1−06
CCF
1
0.8
0.6
0.4
0.2
0
0
4
years
PLEKHA5
KIAA1211
untr
untr2
time point

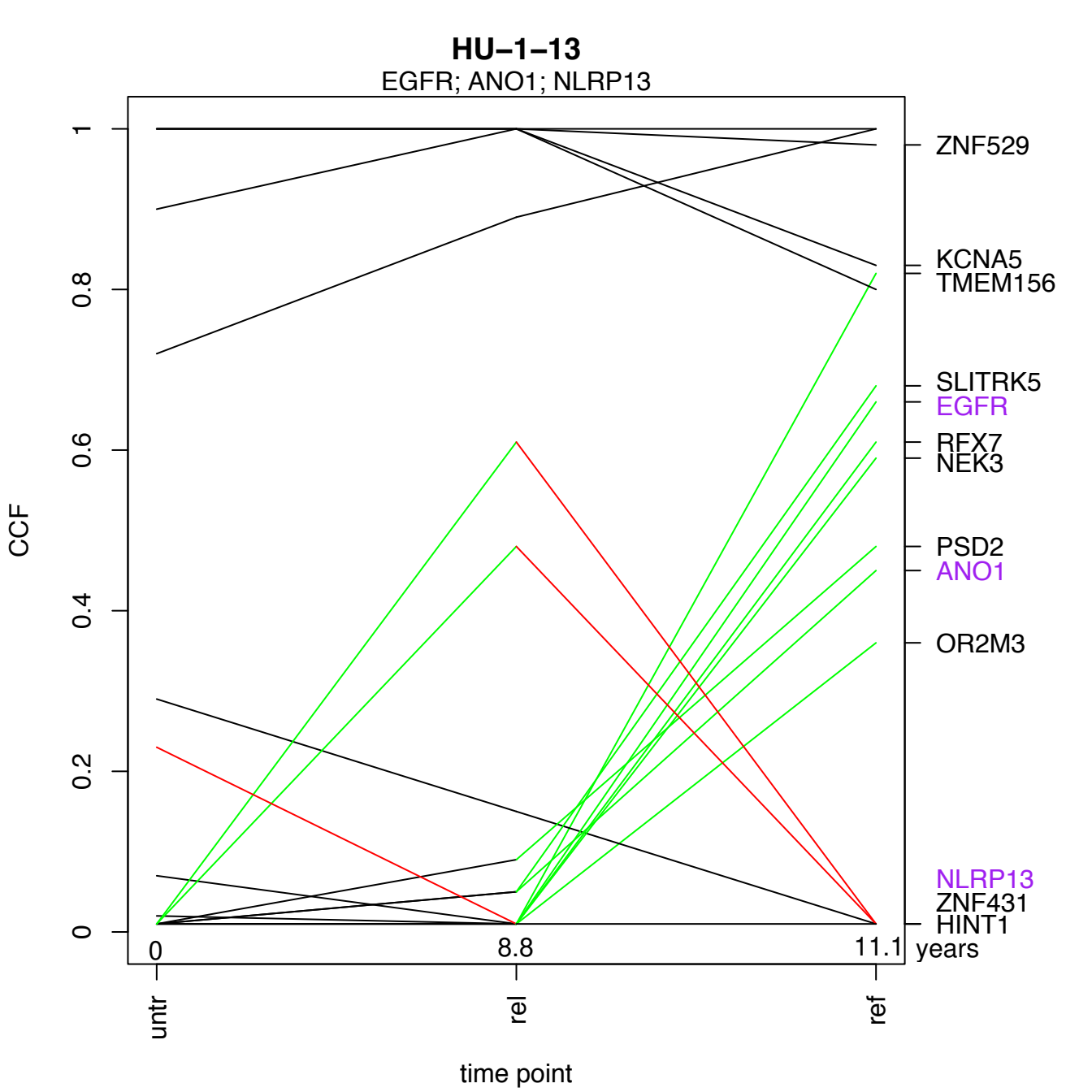

HU−1−07
MYH2
RPS15
CCF
1
0.8
0.6
0.4
0.2
0
0
8.9
years
untr
rel
time point

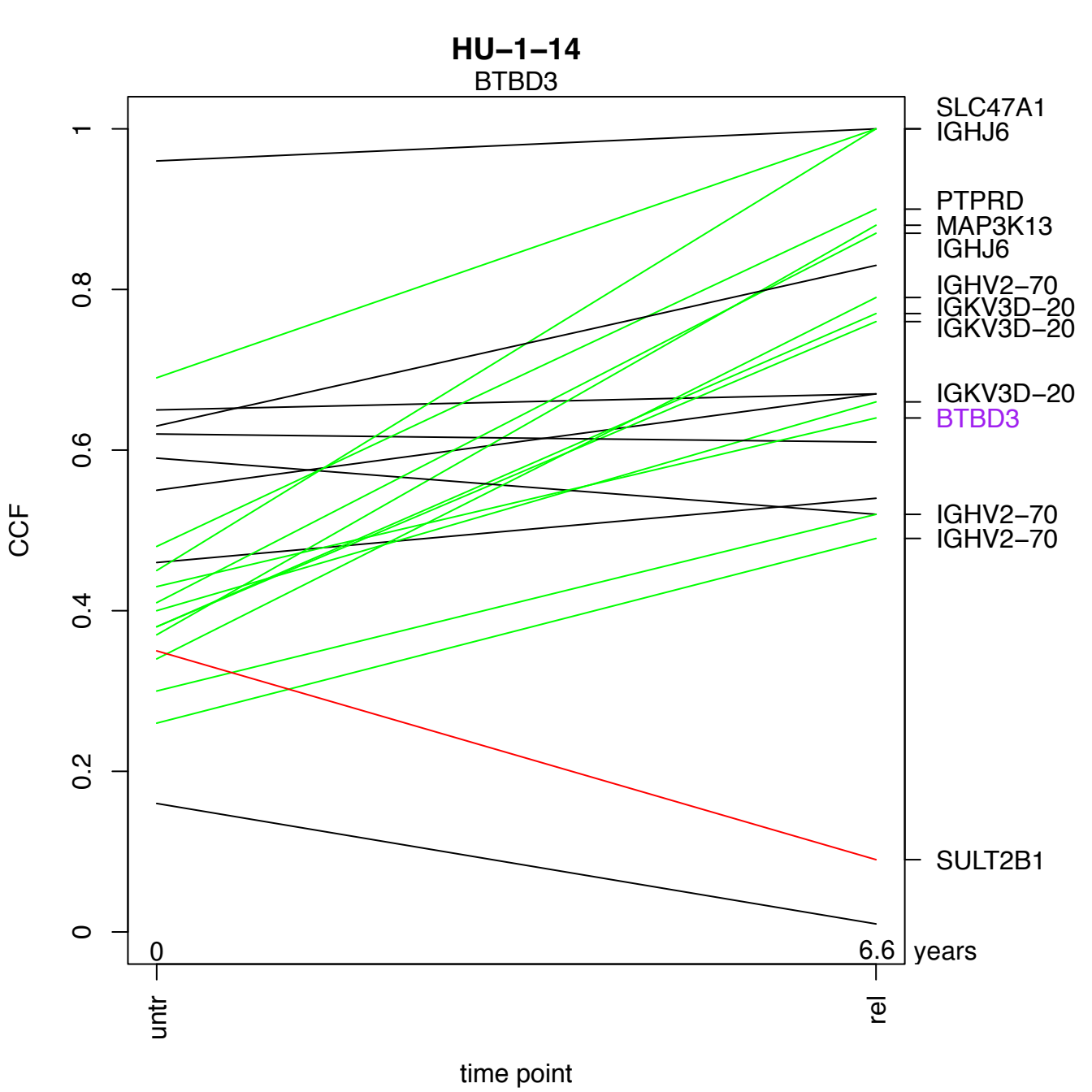

HU−1−08
RIMBP2
PCDH15
CCF
1
0.8
0.6
0.4
0.2
0
0
5.1
years
untr
untr2
time point

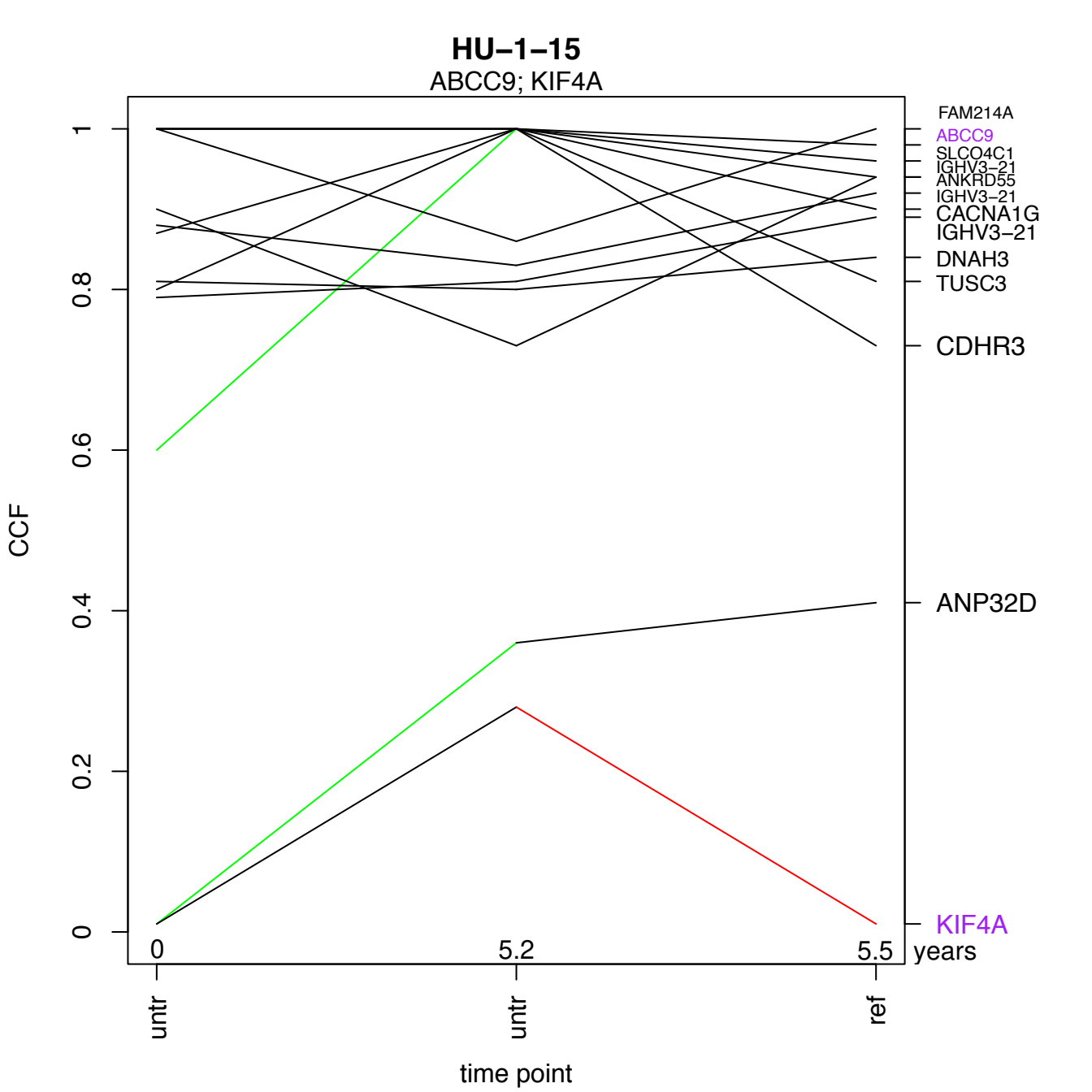

HU−1−09
CCF
1
0.8
0.6
0.4
0.2
0
IGLV1−40
0
8.5
years
untr
untr2
time point

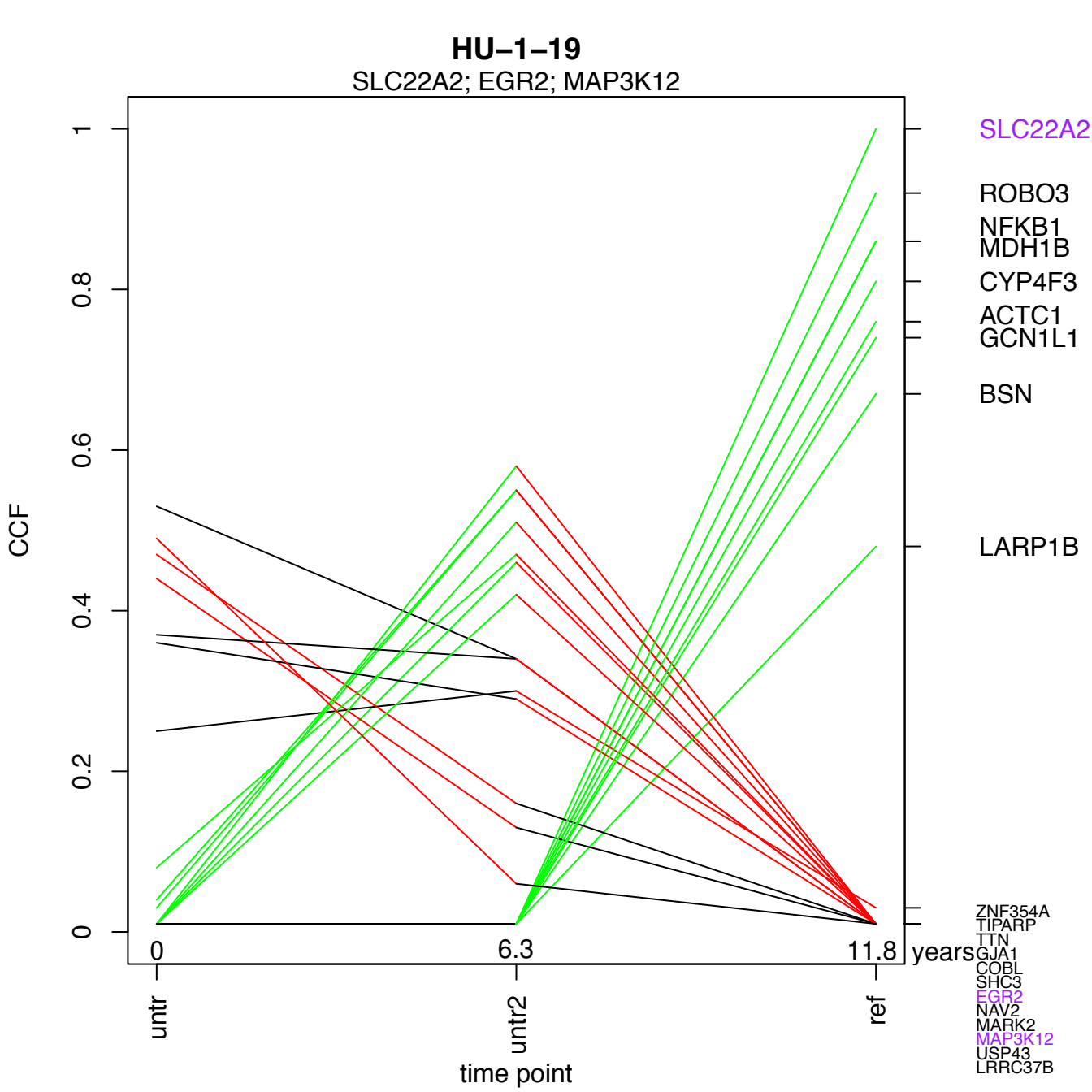

HU−1−10
ATM
APIP
PRRG4
PDHX
PAMR1
CCF
0
0.2
0.4
0.6
0.8
1
0
4.7
years
untr
rel
time point

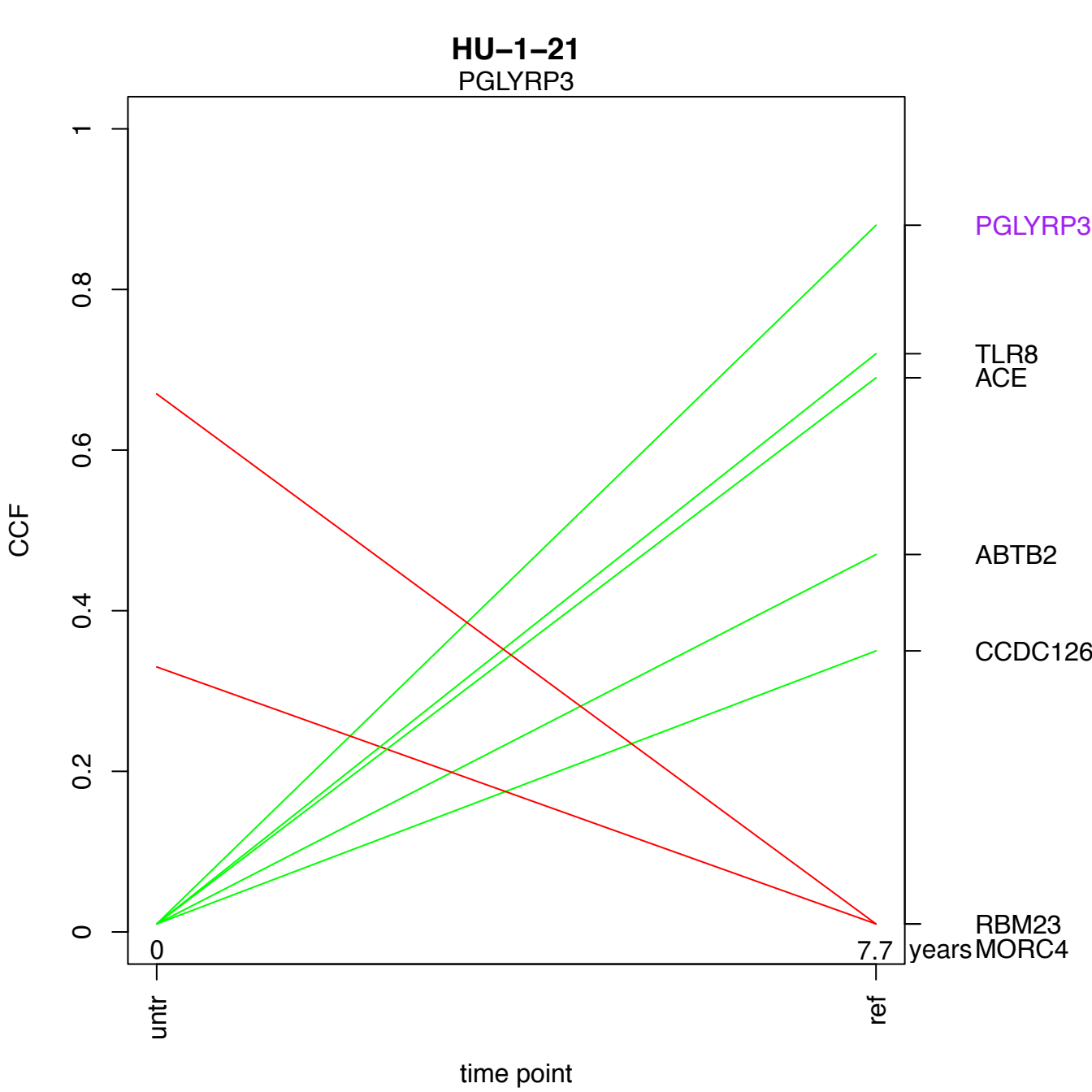

HU−1−11
CCF
0
0.2
0.4
0.6
0.8
1
0
4
years
untr
rel
time point

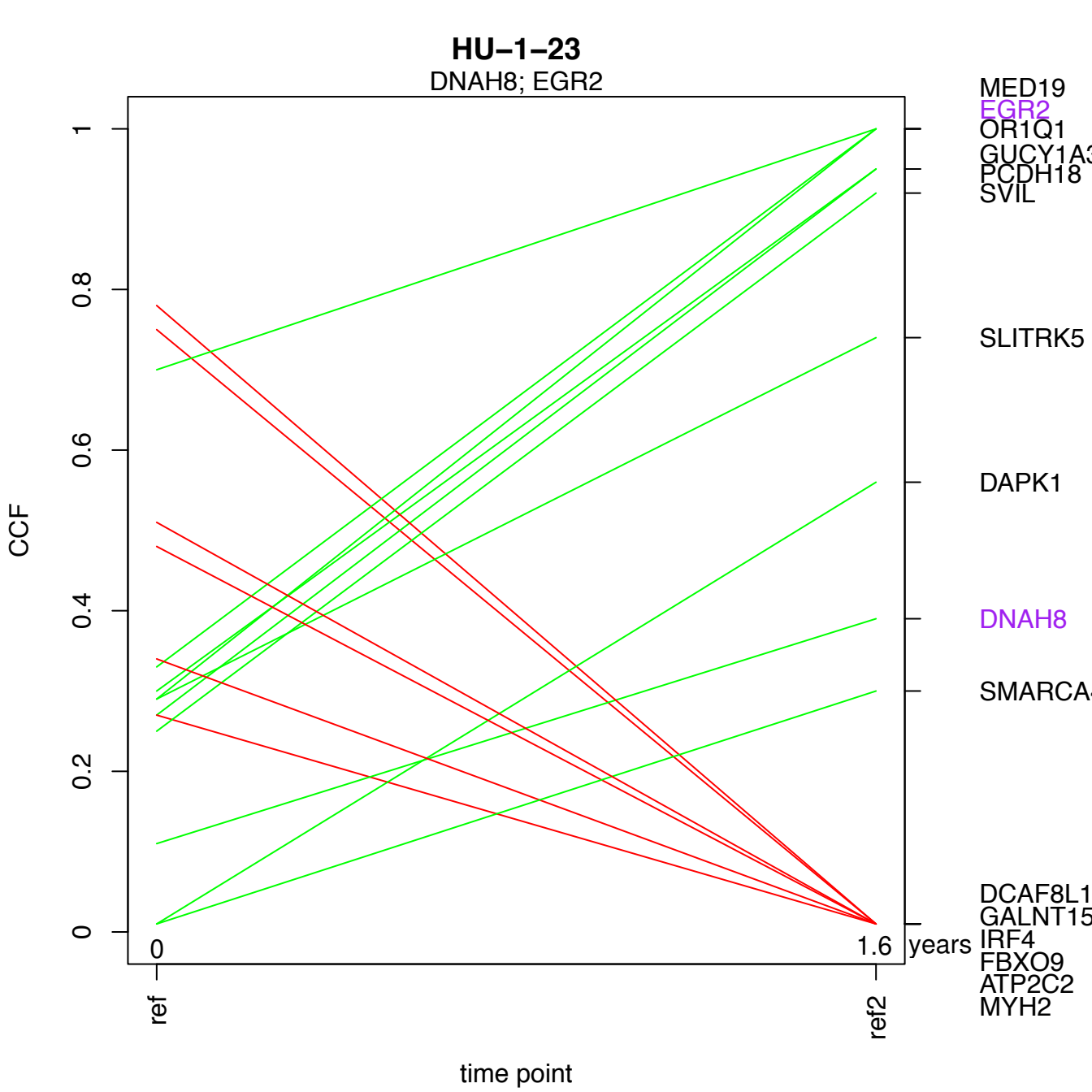

HU−1−13
EGFR; ANO1; NLRP13
CCF
1
0.8
0.6
0.4
0.2
0
0
8.8
11.1
years
untr
rel
ref
time point
ZNF529
KCNA5
TMEM156
SLITRK5
EGFR
RFX7
NEK3
PSD2
ANO1
OR2M3
NLRP13
ZNF431
HINT1


HU−1−14
BTBD3
CCF
0
0.2
0.4
0.6
0.8
1
SLC47A1
IGHJ6
PTPRD
MAP3K13
IGHJ6
IGHV2−70
IGKV3D−20
IGKV3D−20
IGKV3D−20
BTBD3
IGHV2−70
IGHV2−70
SULT2B1
0
6.6
years
untr
rel
time point


HU−1−15
ABCC9; KIF4A
CCF
1
0.8
0.6
0.4
0.2
0
FAM214A
ABCC9
SLCO4C1
IGHV3−21
ANKRD55
IGHV3−21
CACNA1G
IGHV3−21
DNAH3
TUSC3
CDHR3
ANP32D
KIF4A
0
5.2
5.5
years
untr
untr
ref
time point

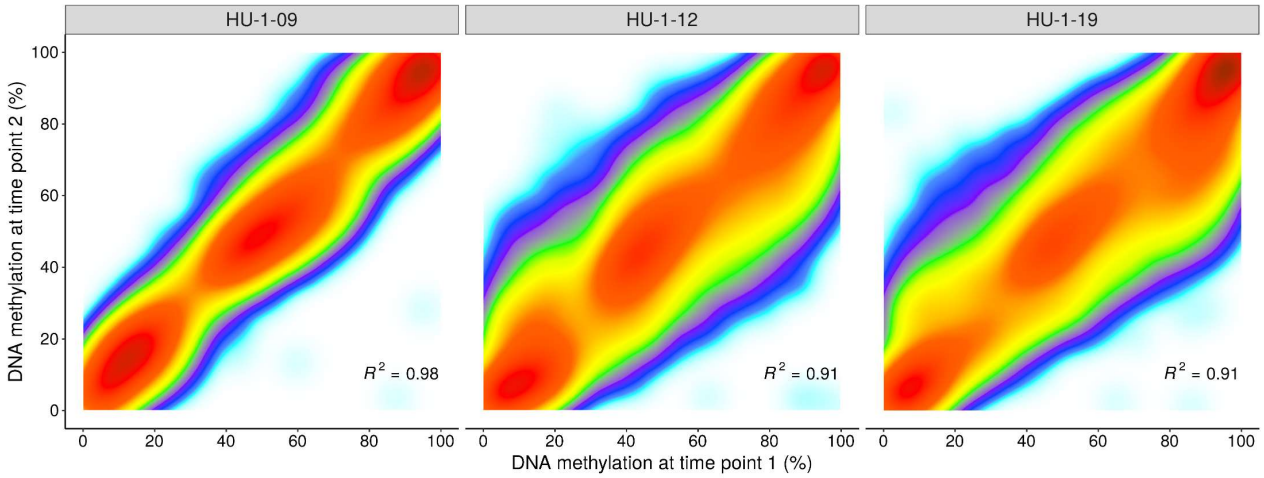

HU−1−19
SLC22A2; EGR2; MAP3K12
CCF
0
0.2
0.4
0.6
0.8
1
0
6.3
11.8
years
untr
untr2
ref
time point
SLC22A2
ROBO3
NFKB1
MDH1B
CYP4F3
ACTC1
GCN1L1
BSN
LARP1B
ZNF354A
TIPARP
TTN
GJA1
COBL
SHC3
EGR2
NAV2
MARK2
MAP3K12
USP43
LRRC37B


HU−1−21
PGLYRP3
CCF
0
0.2
0.4
0.6
0.8
1
0
7.7
years
untr
ref
time point
PGLYRP3
TLR8
ACE
ABTB2
CCDC126
RBM23
MORC4

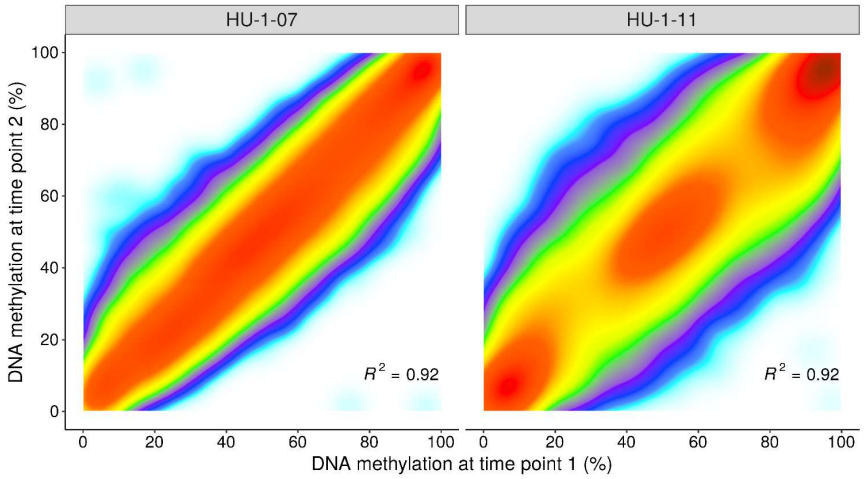

HU−1−23
DNAH8; EGR2
CCF
0
0.2
0.4
0.6
0.8
1
0
1.6
years
ref
ref2
time point
MED19
EGR2
OR1Q1
GUCY1A3
PCDH18
SVIL
SLITRK5
DAPK1
DNAH8
SMARCA
DCAF8L1
GALNT15
IRF4
FBXO9
ATP2C2
MYH2

Suppl. figure 5: Density contour plots of methylation values at two consecutive time points for long-term untreated, relapsed and refractory CLL cases. For each patient, the methylation values of the overall 40,000 most variable CpGs were used to calculate the square of the Pearson correlation coefficient (R^2) between time points.

Long-term untreated phases



Relapsed phases



Refractory phases